

Schriften zur
Informations-
wissenschaft 70

Maria Gäde, Violeta Trkulja, Vivien Petras (Eds.)

**Everything Changes, Everything
Stays the Same? Understanding
Information Spaces**

Proceedings of the 15th International
Symposium of Information Science
(ISI 2017)



Hochschulverband
Informationswissenschaft (HI) e.V.

vwh

Gäde/Trkulja/Petras (Eds.)

Everything Changes, Everything Stays the Same?

**Maria Gäde/Violeta Trkulja/Vivien Petras
(Eds.)**

Everything Changes, Everything Stays the Same? Understanding Information Spaces

Proceedings of the 15th International Symposium
of Information Science (ISI 2017)

Berlin, Germany, 13th–15th March 2017

vwh

Verlag Werner Hülsbusch
Fachverlag für Medientechnik und -wirtschaft

M. Gäde/V. Trkulja/V. Petras (Eds.):

Everything Changes, Everything Stays the Same? Understanding Information Spaces

Bibliografische Information der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter <http://dnb.ddb.de> abrufbar.

Bibliographic information of the German National Library

The German National Library lists this publication in the German National Bibliography; detailed bibliographic data is available online at <http://www.d-nb.de>.

© Verlag Werner Hülsbusch, Glückstadt, 2017

vwh Verlag Werner Hülsbusch
Fachverlag für Medientechnik und -wirtschaft

www.vwh-verlag.de

Einfache Nutzungsrechte liegen beim Verlag Werner Hülsbusch, Glückstadt.

The Verlag Werner Hülsbusch, Glückstadt, owns rights of use for the printed version of this work.

Markenerklärung: Die in diesem Werk wiedergegebenen Gebrauchsnamen, Handelsnamen, Warenzeichen usw. können auch ohne besondere Kennzeichnung geschützte Marken sein und als solche den gesetzlichen Bestimmungen unterliegen.

All trademarks used in this work are the property of their respective owners.

Satz/Typesetting: Werner Hülsbusch, Glückstadt

Druck und Bindung/Printing and binding: SOWA Sp. z o. o., Piaseczno

Printed in Poland

ISSN: 0938-8710

ISBN: 978-3-86488-117-6

Table of Contents

Organization and Sponsors	13
ISI 2017 Programme Committee	15
Editors' Note	17
Keynotes	
<i>Fabian Hemmert</i>	
Information Spaces	20
Building Bridges between the Digital and the Physical World	
<i>Melanie Feinberg</i>	
Translating Texture Data between Information Spaces	21
Session 1: Information Behavior	
<i>Kathrin Rust, Christian Schlögl and Dongback Seo</i>	
Comparing Information Literacy of Students from University of Graz (Austria) and Chungbuk National University (Republic of Korea)	24
<i>Anastasia Christoforidis, Ben Heuwing and Thomas Mandl</i>	
Visualising Topics in Document Collections	37
An Analysis of the Interpretation Processes of Historians	
<i>Barbara Ströhl, Kilian Ganz, Stephanie Richter, Kilian Zieglmeier, Rainer Hammwöhner †</i>	
An Eye-Tracking Study on Differences in Information Transfer by Infographics	50
<i>Dirk Lewandowski, Sebastian Sünkler and Friederike Kerkmann</i>	
Are Ads on Google Search Engine Results Pages Labeled Clearly Enough?	62
The Influence of Knowledge on Search Ads on Users' Selection Behaviour	

Session 2: Designing Scholarly Information Systems

Keith Jeffery, Carlo Meghini, Cesare Concordia, Theodore Patkos, Valerie Brasse, Jacco van Ossenbruck, Yannis Marketakis, Nikos Minadakis and Eda Marchetti

A Reference Architecture for Virtual Research Environments 76

Corinna Breitingner and Bela Gipp

VirtualPatent 89

Enabling the Traceability of Ideas Shared Online
Using Decentralized Trusted Timestamping

Joeran Beel, Siddharth Dinesh, Philipp Mayr, Zeljko Carevic and Raghvendra Jain

Stereotype and Most-Popular Recommendations in the Digital Library
Sowiport 96

Session 3: User Perceptions of Information Systems

Tjaša Jug and Maja Žumer

Content, Physical Appearance, Copy Condition 110

Tagging Customer Book Reviews

Valerie Claessen, Adrian Schmidt and Tamara Heck

Virtual Assistants 116

A Study on the Usability and User Perception of Customer Service
Systems for E-Commerce

Alexander von Janta Lipinski, Henrik Weber, Ralph Koelle and Thomas Mandl

Gamification Elements and Their Perception by Different Gamer Types 131

A Case Study for a Project Management Software

Session 4: Information System Evaluation

Eva-Maria Meier, Patricia Böhm and Christian Wolff

Comparing Heuristic Walkthrough and User Studies in Evaluating
Digital Appliances 146

Wolfgang Semar, Elena Mastrandrea and Fabian Odoni

- Development of a Benchmark System for Social Enterprise Software Benchmark System and Visualisation for Analysing Personal Knowledge Behaviour 158

Juliane Stiller and Péter Király

- Multilinguality of Metadata 164
Measuring the Multilingual Degree of Europeana's Metadata

Anne Chardonrens and Simon Hengchen

- Text Mining for User Query Analysis 177
A 5-Step Method for Cultural Heritage Institutions

Session 5: Metrics and Altmetrics

Edgar Schiebel, Davide Bianchi and András Vernes

- A Bibliometric Framework to Identify and Delineate Subfields of Research on Tribological Wear 192
Part One: Can We Identify Fundamental Issues by Clusters of Similar Journals?

Elisabeth Vogler, Christoph Schindler, Alexander Botte and Marc Rittberger

- Are Altmetrics Effective in Transdisciplinary Research Fields? 203
Altmetric Coverage of Outputs in Educational Research

Session 6: (Social) Media Analysis

Felix Hamborg, Norman Meuschke, Corinna Breitingner and Bela Gipp

- news-please* 218
A Generic News Crawler and Extractor

Felix Hamborg, Norman Meuschke, Akiko Aizawa and Bela Gipp

- Identification and Analysis of Media Bias in News Articles 224
The Impact of Objectivity and Believability on Corporate Decision Making and Performance

Florian Meier, Johannes Aigner and David Elswailer

- Using Sessions from Clickstream Data Analysis to Uncover
Different Types of Twitter Behaviour 237

Panels

*Wolf Rauch, Rainer Kuhlen, Wolfgang G. Stock, Christian Wolff,
Christa Womser-Hacker and Christian Schlögl*

- Significance and Relevance of Information Science in
German-language Countries 252

A Panel Discussion Devoted to the 65th Birthday of Wolf Rauch

Antje Michel

- “Information and Learning” 264

Research at the Interface between Information Science and the
Learning Sciences

Michael Seadle

- Examining Research Integrity 269

Gerhard Lustig Award Papers

Fabian Muster

- Data Strategy Model 274

A Reference Model to Develop Data Strategies

Roland List

- Relationship between Downloads and Citations 278

The Case of Two Economics Journals

Alexandra Linhart

- Google Now and User Data 281

A Diary Study on Perceptions of Collection and Use of Personal Data
by Google Now Cards

Helene Brinken and Helena Kock

Exploring Intergenerational Knowledge Sharing in Organizations 285
Potentials and Limitations of Information and Communication Technology
(ICT)

Sandra Balck

The X(Disciplinarity) of Information Science 290

Robert Jackermeier

Interactive Pedestrian Indoor Localization 293

Doctoral Colloquium I

Information Literacy and Behavior, Knowledge Management, Visualization

Stefan Dreisiebner

Development and Evaluation of a MOOC Extension for Teaching Sub-
ject-Specific Skills in Information Literacy for Business Administration 298

Stefanie Havelka

Mobile Privacy and Apps: Investigating User Behavior and Attitude 303

Julian Hocker

Visualization for Text Mining in the Digital Humanities 308
Empowering Researchers to Use Advanced Tools for Text Mining

Fabian Odoni

Development of Knowledge Management Performance Metrics
for Enterprise Social Software 314

Veit Stephan

Information Visualization of Environment Maps for Complex, Relational
Data for Better Reflected Decision Making on the Example of Food 321

Doctoral Colloquium II

Information Retrieval, Library Information Science, Social Media

Katrin Behnert

Relevance Clues 330

Developing an Experimental Design to Examine the Criteria Behind
Relevance Judgments

Anna Klamet

E-Publishing in Small Nations of the European Union 336

A Comparative Approach to Policy Frameworks and SME Business
Strategies

Anthea Moravánszky

How Can Social Software Enhance Training Measures within
Software Implementations in the Financial Sector? 343

Kaltrina Nuredini

Enriching a Library Portal (*EconBiz*) with Altmetrics 348

Poster Presentations

*Manuel Burghardt, Katharina Hafner, Laura Edel, Sabrin-Leila Kenaan
and Christian Wolff*

An Information System for the Analysis of Color Distributions
in *MovieBarcodes* 356

Nicolas Prongué and René Schneider

Data Streams in *linked.swissbib.ch* 359
The Swiss Metacatalog in the Linked Open Data Cloud

Achim Oßwald

For You, Me, Them? 362
Choosing Designated Communities for Personal Digital Archiving

*Christoph Schindler, Julian Hocker, Lars Müller, Maria Maleshkova
and Tobias Weller*

Grasping the Materializations of Practices in Digital Humanities 365
A Semantic Research Environment for Analyzing Exam Grading Practices
in German High Schools

Stefan Dreisiebner

Information Literacy Online 368
An Erasmus+ Project to Improve Students' Competencies

Vera Hillebrand and Elke Greifeneder

International Degree Mobility in Library and Information Science 371

Peter Böhm, Christoph Schindler, Heiko Rölke and Marc Rittberger

Limits and Potentials of Grasping User Interaction in Controlled and
Open World Environments 374
Contrasting PIAAC Large-Scale-Assessment Usage Data with
Web Analytics

Thomas Mandl and Tatjana Aparac Jelušić

Lowering the Barriers to Entry into LIS Master Programs 377
The EINFOSE Project

Nicolas Apostolopoulos, Christoph Schimmel and Ilker Egilmez

Open Encyclopedia System 380
Open Source Software for Open Access Encyclopedias

Fabian Odoni, Wolfgang Semar and Elena Mastrandrea

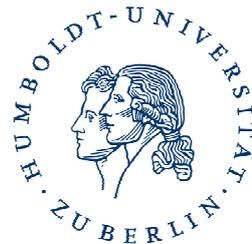
VICI 386
Visualisation of Collaboration in Social Enterprise Software Systems

Simone Kronenwett

Research Data at the Faculty of Arts and Humanities of the University
of Cologne 389

Organization and Sponsors

Berlin School of Library and Information Science
Humboldt-Universität zu Berlin
www.ibi.hu-berlin.de



Hochschulverband Informationswissenschaft (HI) e.V.
www.informationswissenschaft.org



DIPF Education Research and Educational Information
www.dipf.de



Programme Chairs & Local Management

Maria Gäde, Humboldt-Universität zu Berlin
Vivien Petras, Humboldt-Universität zu Berlin
Violeta Trkulja, Humboldt-Universität zu Berlin

Paper Chair

Vivien Petras, Humboldt-Universität zu Berlin

Poster Chair

Marc Rittberger, DIPF

Panel Chair

Hans-Christoph Hobohm, University of Applied Sciences Potsdam

Gerhard Lustig Award Chair

Achim Oßwald, Cologne University of Applied Sciences

Student Programme Chairs

Maria Fentz, Humboldt-Universität zu Berlin

Hanna Kummel, Chur University of Applied Sciences

Doctoral Colloquium Chair

Christian Wolff, Regensburg University

Sponsors

Inficon

<http://www.inficon.com/>



Facet Publishing

<http://www.facetpublishing.co.uk>



ISI 2017 Programme Committee

Tatjana Aparac Jelusic, University of Zadar

Christoph Bläsi, Study of the Book / Johannes Gutenberg University of Mainz

Stephan Büttner, University of Applied Sciences Potsdam

Manuel Burghardt, Regensburg University, Media Computing

Ernesto William De Luca, Georg-Eckert-Institute – Leibniz-Institute for International Textbook Research

Maximilian Eibl, Chemnitz Technical University

David Elsweiler, Regensburg University

Norbert Fuhr, University of Duisburg-Essen

Maria Gäde, Humboldt-Universität zu Berlin, Berlin School of Library and Information Science

Ursula Georgy, Cologne University of Applied Sciences

Juan Gorraiz, University of Vienna

Joachim Griesbaum, University of Hildesheim

Vicente Pablo Guerrero-Bote, University of Extremadura

Siegfried Handschuh, University of Passau

Stefanie Haustein, University of Montréal

Gerhard Heyer, Leipzig University, Institute of Computer Science

Hans-Christoph Hobohm, University of Applied Sciences Potsdam

Isto Huvila, Uppsala University

Dirk Lewandowski, Hamburg University of Applied Sciences

Bernd Ludwig, Regensburg University

Thomas Mandl, University of Hildesheim

Zinaida Manžuch, Vilnius University

Philipp Mayr, GESIS – Leibniz Institute for the Social Sciences

Michel Menou, University College London, School of Library, Archives and Information Services

Antje Michel, University of Applied Sciences Potsdam

Heike Neuroth, University of Applied Sciences Potsdam

Achim Oßwald, Cologne University of Applied Sciences

Franjo Pehar, University of Zadar, Department of Information Sciences

Isabella Peters, ZBW – Leibniz Information Centre for Information Economics

Vivien Petras, Humboldt-Universität zu Berlin, Berlin School of Library and Information Science

Wolf Rauch, University of Graz

Gerhard Reichmann, University of Graz

Ulrich Reimer, University of Applied Sciences St. Gallen

Marc Rittberger, DIPF – Educational Research and Educational Information

Tefko Saracevic, Rutgers University

Christian Schlögl, University of Graz, Institute of Information Science and Information Systems

Jasmin Schmitz, ZB MED – Leibniz Information Centre for Life Sciences

Rene Schneider, Haute Ecole de Gestion, Geneva

Michael Seadle, Humboldt-Universität zu Berlin, Berlin School of Library and Information Science

Diane Sonnenwald, University of Copenhagen

Elisabeth Staudegger, University of Graz

Juliane Stiller, Humboldt-Universität zu Berlin, Berlin School of Library and Information Science

Wolf Stock, Heinrich Heine University Düsseldorf, Department of Information Science

Mike Thelwall, University of Wolverhampton

Violeta Trkulja, Humboldt-Universität zu Berlin, Berlin School of Library and Information Science

Sirje Virkus, Tallinn University

Erich Weichselgartner, ZPID – Leibniz Institute for Psychology Information

Katrin Weller, GESIS – Leibniz Institute for the Social Sciences

Christian Wolff, Regensburg University, Media Computing

Dietmar Wolfram, University of Wisconsin-Milwaukee

Christa Womser-Hacker, University of Hildesheim

Maja Žumer, University of Ljubljana

Editors' Note

The 15th International Symposium of Information Science, a biannual conference for scholars, researchers, professionals and students in information science, will be held at Humboldt-Universität zu Berlin from 13 to 15 March 2017. This year's theme of "Everything Changes, Everything Stays the Same? Understanding Information Spaces" puts the dynamic of the field of information and its applications in the center.

The conference call explicates the theme: "With the ongoing digitization and virtualization of goods, services and living environments, information science reflects on the potential changes within the information society. While some parts of the community proclaim a revolutionary shift not only in the way we approach information and information systems, but also in the way society constitutes itself, others state that while the applications and interfaces adapt to advances in information technology, the underlying principles for human interactions with information remain the same."

At ISI 2017, contributions on the development of innovative information spaces and services, on analyses of human-computer interactions in physical or virtual information spaces and on the role of information science and its research endeavors were invited.

The two keynotes by Fabian Hemmert and Melanie Feinberg zoom in on various aspects of the theme, focusing on the transition between the physical and digital world and on the changes in meaning when objects transition between information spaces, respectively.

The submitted papers, posters and panel contributions were reviewed in a double blind peer review process. The programme committee consisted of 61 members from 13 countries. Of the 48 submissions (32 papers, 13 posters, 3 panels), the programme committee selected 15 long papers, 4 short papers, 11 posters and all 3 panels, resulting in a paper acceptance rate of 59%.

The conference is organized into two parallel tracks, presenting short and long papers in 6 paper sessions showcasing the depth and breadth of information science: information behavior, designing scholarly information systems, user perceptions of information systems, information systems evaluation,

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 17–18.

metrics & altmetrics, and (social) media analysis. The paper sessions are complemented by 3 panels (significance and relevance of information science in German-language countries, “information and learning” – research at the interface between information science and the learning sciences, and examining research integrity), and a poster session.

A particular focus of ISI conferences has always been the presentation of student and doctoral work. The Gerhard Lustig Award is awarded to the best Master thesis in information science of the last two years before the conference. One conference session is devoted to 6 Master thesis presentations, from which the award winner will be selected. The student session provides a safe space for students to discuss their work, while the doctoral colloquium presents a possibility to share research ideas with a larger mentoring group.

A satellite workshop on the relationship of information science and the digital humanities builds further bridges with other disciplines.

A conference can only be organized with the help of many people. We would like to thank the contributors to the research programme for submitting stimulating and innovative contributions and the members of the programme committee for their timely and constructive reviews. A special thank you to our partners, the German Association for Information Science (HI) and DIPF (Educational Research and Educational Information) for providing substantial support to the conference. Without their contributions, a successful conference would not have been possible. A special thank you to the conference sponsors, especially for the support of young scholars. We would also like to thank the voluntary helpers, mostly students, from Humboldt-Universität zu Berlin and DIPF, who ensure a streamlined local organization. The publisher Werner Hülsbusch once again provided the proceedings both in print and in an Open Access version in a timely manner – we are very grateful for the productive collaboration.

We wish all participants a successful, productive and inspiring conference.

Maria Gäde, Vivien Petras, Violeta Trkulja

Keynotes

Information Spaces

Building Bridges between the Digital and the Physical World

Fabian Hemmert

University of Wuppertal, Germany
fabian@hemmert.eu

Abstract

As humans, we are used to being, living and thinking in the three-dimensional, physical space that is called the “world”. It is our natural habitat.

Recently, we have become active in another world: the “digital world”, as it is often called. Unfortunately, most of our interactions with the digital world are isolated from the world that is our natural habitat. These interactions happen on screens of various sizes, either hand-held or fixed in front of our eyes (in bulky helmets or “smart” glasses). Sadly, technology is getting more and more between us.

In his keynote, Prof. Dr.-Ing. Fabian Hemmert will show us some alternatives: his prototypes explore ways of making digital information tangible, placing it in our hands and in our world. He will show us shape-changing devices and interactive rooms that interweave information and physical space. His talk will be a guided tour through a possible future of human-computer interaction, in which digital information can be retrieved intuitively, in computationally-enriched interactions with the natural world.

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, p. 20.

Translating Texture

Data between Information Spaces

Melanie Feinberg

School of Information and Library Science
University of North Carolina, Chapel Hill, U.S.
melanie_feinberg@unc.edu

Abstract

In information studies, we tend to think that data retains the same meaning as it moves from one information space to another. When a system changes its interface, or when data moves from one system to another, the data itself doesn't change. A library catalog record doesn't change when the catalogue interface changes, for example. Or does it?

In this keynote, I discuss how information spaces contribute to information meaning. I describe a project to translate some purposefully weird, experimental information collections from one kind of database implementation (a relational-style database) to another kind of database implementation (a graph-style database). I focus this translation project on the conceptual lens of texture: the relationship between elements in a composition. Describing the texture of a musical piece, for example, is a way of talking about how melody, harmony, and rhythm combine to produce a particular quality of sound. Here, I use texture to talk about how data and space combine to produce a particular quality of meaning. I ask: how can texture be maintained when data is moved from one information space to another?

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, p. 21.

Session 1:
Information Behavior

Comparing Information Literacy of Students from University of Graz (Austria) and Chungbuk National University (Republic of Korea)

Kathrin Rust¹, Christian Schlögl¹, Seo DongBack²

¹ University of Graz, Austria

{kathrin.rust, christian.schloegl}@edu.uni-graz.at

² Chungbuk National University, Republic of Korea

dseo@chungbuk.ac.kr

Abstract

In this article, we present the results of a study in which we compared information literacy of master students from the University of Graz, Austria and Chungbuk National University (CBNU), Republic of Korea. Data were collected using a multiple-choice questionnaire which consisted of the following parts: demographic data, self-assessment of one's information literacy, use of information sources, and knowledge test. The latter was designed on the basis of the Information Literacy Competency Standards for Higher Education (ALA, 2000). Data were collected in two classes of the Business Administration program at the University of Graz. At CBNU, the study participants were subscribed to the master programs of Business Administration, Management Information Systems, and Psychology. Usually, it took the students 20 minutes to fill out the questionnaire.

The results reveal that students from the University of Graz have a higher level of information literacy than their colleagues from CBNU. To some degree, this might be due to cultural and social differences between the students. However, the test instrument might have had also some cultural (European) bias. Finally, it turned out that the Korean students were more exposed toward the use of modern information and communication technologies.

Keywords: information literacy; comparison; questionnaire; University of Graz; Chungbuk National University; cultural differences

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 24–36.

1 Introduction

In today's information society, skills like searching, assessing, organizing and using information are more important than ever. This is not only true in the leisure time and at work but in particular in higher education. Information literacy has become a key qualification nowadays. So far, there exist several definitions on information literacy. According to the American Library Association (1989), information literacy is a set of abilities which are required to "recognize when information is needed and ... to locate, evaluate, and use effectively the needed information."

Meanwhile several models and standards have been developed which try to shed more light on this concept. While models try to explain the entire information process or parts of it, like for instance the "Big 6 Skills" by Eisenberg and Berkowitz (1990) or the "Information Search Process" by Kuhlthau (1991), standards concentrate more on the skills information literacy consists of. Probably the best known of them are the Information Literacy Competency Standards for Higher Education by the American Library Association (2000). In the meantime, the Association of College and Research Libraries (2016) has provided an information literacy "framework" which tries to provide more "flexible options for implementation, rather than ... a set of standards or learning outcomes, or any prescriptive enumeration of skills."

Many information literacy studies have been performed in particular in Anglo-American countries so far. In contrast, information literacy is not so well investigated in German-language countries (for instance, Beutelspacher, 2014, or Klatt et al., 2001). Also the authors of this paper have already conducted two similar studies in which they investigated information literacy on the basis of a questionnaire/knowledge test. In the first study, Beutelspacher, Henkel and Schlögl (2015) demonstrated that students in the bachelor program on business administration could clearly improve their level of information literacy in a specific course devoted to this topic. The second study analyzed the level of information literacy of student beginners from six bachelor programs each of which was offered by different faculties at the University of Graz. The results revealed that student beginners had only moderate information competencies which varied more or less strongly between the six bachelor programs/faculties (Maurer, Schlögl & Dreisiebner, 2017). How-

ever, to our knowledge no information literacy study which focused on cultural differences was conducted so far.

2 Research questions and methodology

In this study we aim at comparing information literacy of master students from the University of Graz and from Chungbuk National University (CBNU). The underlying research questions are:

1. How information literate are master students from the University of Graz and from Chungbuk National University?
2. Are the master students from the University of Graz more information literate than those from CBNU?
3. How do the students assess their own level of information literacy in comparison to their actual information literacy skills?

To avoid that the test participants require subject specific knowledge, we used the same knowledge test as in the previous study (study beginners) as a starting point (cf. Maurer, 2016). However, it turned out after the pre-test that a few questions had to be slightly changed in the Korean language version of the questionnaire. No changes were necessary in the German language version (for the master students from Graz).¹

Like in the previous studies, the questionnaire consisted of four parts:

- personal/demographic data
- self-assessment of the level of information literacy by students
- knowledge test
- knowledge about use of information sources for study purposes.

Data collection at the University of Graz took place in the two Business Administration master courses “Retail Marketing” and “Seminar on Information Science and Information Systems” in May 2016. In total, 41 business administration students filled out the questionnaire.

Since it is common to start a job after the end of the bachelor program in Republic of Korea, it turned out to be more difficult to get a similar number of questionnaires at CBNU. Therefore, we decided to include also the master

¹ The two questionnaires are included in Rust (2016: 74–90).

programs of Management Information Systems and of Psychology besides Business Administration in the survey. At CBNU, data collection took place in the calendar weeks 19–22. Altogether, we received 39 questionnaires from CBNU students. Usually, it took the Austrian and Korean students 20 minutes to complete the questionnaire. In seldom cases, when students had problems to understand a question, they had the possibility to ask for clarification.

3 Results

In this section, we follow the structure of the questionnaire. First we describe the demographic data of the survey participants from the University of Graz and from CBNU. Afterwards, we present the results of the knowledge test and compare the outcomes of the students from the University of Graz and CBNU. Then, we compare the results of the knowledge test to the self-evaluations of the students. Finally, we provide an insight to the information sources and tools used by the students for their studies and differences between the two universities.

3.1 Demographic data

Tables 1 and 2 show the demographic data of the Austrian and Korean students. Since Business Education students attend partly the same master courses as Business Administration students, the population of the University of Graz consists of these two groups: seven students from Business Education and 34 from Business Administration. Three quarters of the survey participants were female, one quarter was male. The gender distribution is more balanced for students from CBNU (20 female and 19 male). Furthermore, the mean age of CBNU students is slightly higher (26.7 vs. 24.9 years) than that of the students at University of Graz.

Table 1: Gender and age of respondents from the University of Graz

Study	<i>n</i>	Male	Female	Mean age
Business Administration	34	8	26	24.9
Business Education	7	3	4	25.1
Total	41	11	30	24.9

Table 2: Gender and age of respondents from Chungbuk National University

Study	<i>n</i>	Male	Female	Mean age
Management Information Systems	21	11	10	25.9
Business Administration	8	3	5	28.1
Psychology	10	5	5	27.3
Total	39	19	20	26.7

3.2 Results of knowledge test

The knowledge test consisted of 27 single and multiple-choice questions. The correct response to a single-choice question was rewarded with one point, for multiple-choice questions it was possible to get up to 1.5 points. In total, the test participants could receive 31.5 points. For the following analyses, a grade ranging from Very good (1) to Not sufficient (5) was derived on the basis of the total points received (table 3).

Table 3: Knowledge test – grading system

Points	Grade
28–31.5	1 – Very good
24.5–27.5	2 – Good
20–24	3 – Satisfactory
16–19.5	4 – Sufficient
0–15.5	5 – Not sufficient

As was not expected and as can be seen in figure 1, the knowledge test revealed big differences in the information literacy grades between the Austrian and Korean students. 80 percent of the students from the University of Graz were graded either Very good (1) or Good (2). This is in contrast to the students from CBNU, where information literacy of half of the test participants was evaluated Satisfactory (3). 41 percent were graded Sufficient (4) or even Not sufficient (5).

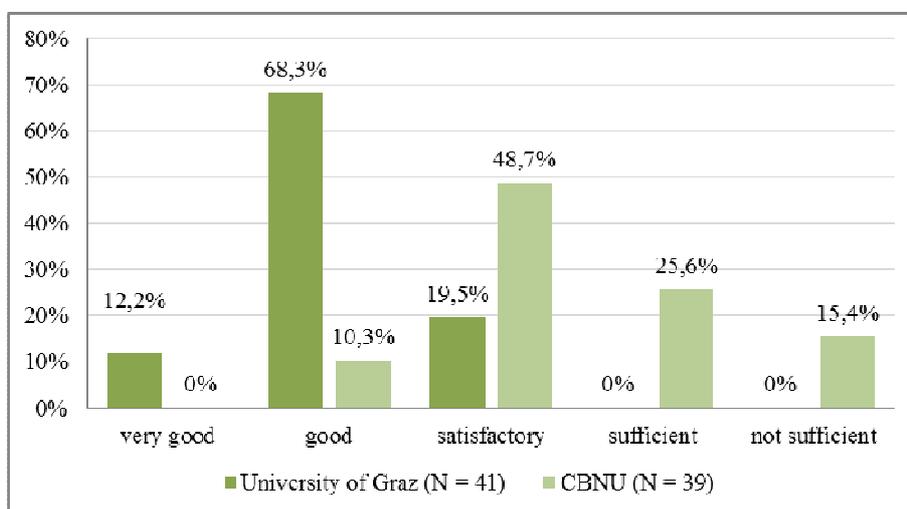


Fig. 1 Overall results of knowledge test: University of Graz vs. Chungbuk National University

3.3 Self-assessment of students

The operationalization of the questionnaire was conducted on the basis of the Information Literacy Competency Standards for Higher Education which distinguishes between five standards (American Library Association, 2000). Since standard 4 deals with the use of information which is more or less domain and subject specific, it was not considered in our survey. In contrast, we considered two aspects of standard 1 when it came to the self-assessment of information literacy by the students: the ability to identify one’s information need and overview about relevant information sources (table 4).

Table 4: Self-assessment statements in the questionnaire

Standard	Statement
Standard 1	Ability to identify one’s information need
Standard 1	Having an overview of relevant information sources
Standard 2	Searching for information on the internet and in databases
Standard 3	Evaluation and quality assessment of information and information sources
Standard 5	Knowledge and observance of ethical, legal and social principles in the handling of information

Table 5: Self-assessment by students from the University of Graz

Grade	Self-assessment				
	Standard 1		Standard 2	Standard 3	Standard 5
	Statement 1	Statement 2			
Very good	10 (24.4%)	10 (24.4%)	19 (46.3%)	7 (17.1%)	15 (36.6%)
Good	28 (68.3%)	26 (63.4%)	22 (53.7%)	15 (36.6%)	17 (41.5%)
Satisfactory	3 (7.3%)	4 (9.8%)	0 (0.0%)	17 (41.5%)	9 (22.0%)
Sufficient	0 (0.0%)	1 (2.4%)	0 (0.0%)	2 (4.9%)	0 (0.0%)
Not sufficient	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Total	41 (100%)	41 (100%)	41 (100%)	41 (100%)	41 (100%)

When comparing tables 5 and 6, it can be seen that students from the University of Graz assessed their level of information literacy higher than students from CBNU with regards to standards 1, 2 and 5. At least one quarter, and in case of standard 2 nearly half of the students from the University of Graz evaluated their competencies Very good (1). The difference in the self-assessment was closer for standard 3 (search for information in the internet and in databases).

Table 6: Self-assessment by students from Chungbuk National University

Grade	Self-assessment				
	Standard 1		Standard 2	Standard 3	Standard 5
	Statement 1	Statement 2			
Very good	5 (12.8%)	2 (5.1%)	5 (12.8%)	1 (2.6%)	1 (2.6%)
Good	11 (28.2%)	17 (43.6%)	20 (51.3%)	15 (38.5%)	14 (35.9%)
Satisfactory	17 (43.6%)	14 (35.9%)	9 (23.1%)	13 (33.3%)	15 (38.5%)
Sufficient	4 (10.3%)	4 (10.3%)	5 (12.8%)	6 (15.4%)	5 (12.8%)
Not sufficient	2 (5.1%)	2 (5.1%)	0 (0.0%)	4 (10.3%)	4 (10.3%)
Total	39 (100%)	39 (100%)	39 (100%)	39 (100%)	39 (100%)

Since each of the questions of the knowledge test can be assigned to a particular standard, it is also possible to compare the self-assessments for each standard with the actual test results (cf. table 7). One interesting result of this comparison is that nearly all students (ca. 85 percent) over-estimate their abilities to identify their information needs and to have an overview about the relevant information sources. This divergence is slightly lower with

regard to standard 2 (information search) with an “over-assessment” by approximately 60 percent of the students.

Contrary to the first two standards, the self-evaluation is more balanced when it comes to the knowledge of ethical, legal and social principles in the handling of information (standard 5). Approximately one third of the students overly estimate, realistically assess, and under-assess their information competencies in comparison with the results of their knowledge tests. In addition, there is little difference between the Austrian and the Korean students in this standard. Interestingly, there is also one standard for which many students under-estimate their competences: the evaluation and quality assessment of information and information sources (standard 3).

These results clearly show that information literacy is a very tricky issue: students usually believe that they are more information literate than they really are. This is particularly true for informational activities they perform very often.

Table 7: Comparison between self-assessment and actual level of information literacy: students from the University of Graz vs. Chungbuk National University

Self-assessment versus knowledge test			
	Realistic	Over	Under
Standard 1			
UG	4 (9.7%)	35 (85.4%)	2 (4.9%)
CBNU	4 (10.3%)	33 (84.6%)	2 (5.1%)
Total	8 (10.0%)	68 (85.0%)	4 (5.0%)
Standard 2			
UG	12 (29.2%)	22 (53.7%)	7 (17.1%)
CBNU	6 (15.4%)	25 (64.1%)	8 (20.5%)
Total	18 (22.5%)	47 (58.8%)	15 (18.7%)
Standard 3			
UG	9 (22.0%)	8 (19.5%)	24 (58.5%)
CBNU	8 (20.5%)	15 (38.5%)	16 (41.0%)
Total	17 (21.3%)	23 (28.7%)	40 (50.0%)
Standard 5			
UG	15 (36.6%)	13 (31.7%)	13 (31.7%)
CBNU	11 (28.2%)	15 (38.5%)	13 (33.3%)
Total	26 (32.5%)	28 (35.0%)	26 (32.5%)

3.4 Knowledge and use of information sources

The usage rates of web search engines and information sources show again differences between the Koreans and Austrians. Nearly all CBNU students (95 percent) use web search engines several times a day. This is true for 78 percent of the students from Graz (table 8).

Table 8: Frequency of use of web search engines

University	Usage of web search engines					Total
	Several times a day	Once a day	Several times a week	Once a week	Never	
UG	32 (78.0%)	4 (9.8%)	5 (12.2%)	0 (0.0%)	0 (0%)	41 (100%)
CBNU	37 (94.9%)	2 (5.1%)	0 (0.0%)	0 (0.0%)	0 (0%)	39 (100%)
Total	69 (86.3%)	6 (7.5%)	5 (6.2%)	0 (0.0%)	0 (0%)	80 (100%)

The analysis concerning the use of information sources and tools for study purposes (fig. 2 and 3) shows that the students from Graz use “traditional” sources like books, online books and library catalogues more often (fig. 2).

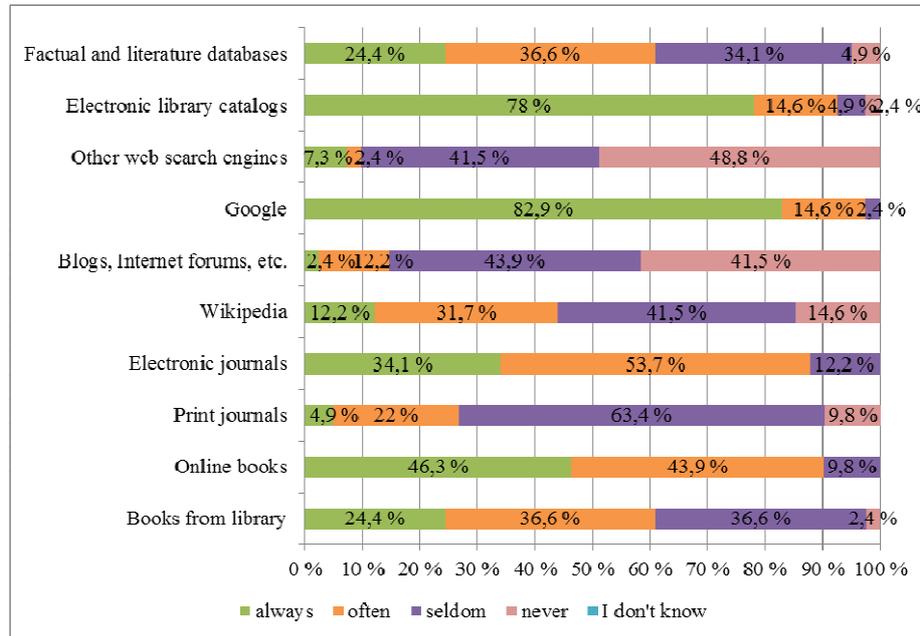


Fig. 2 Use of different information sources and tools by students from the University of Graz

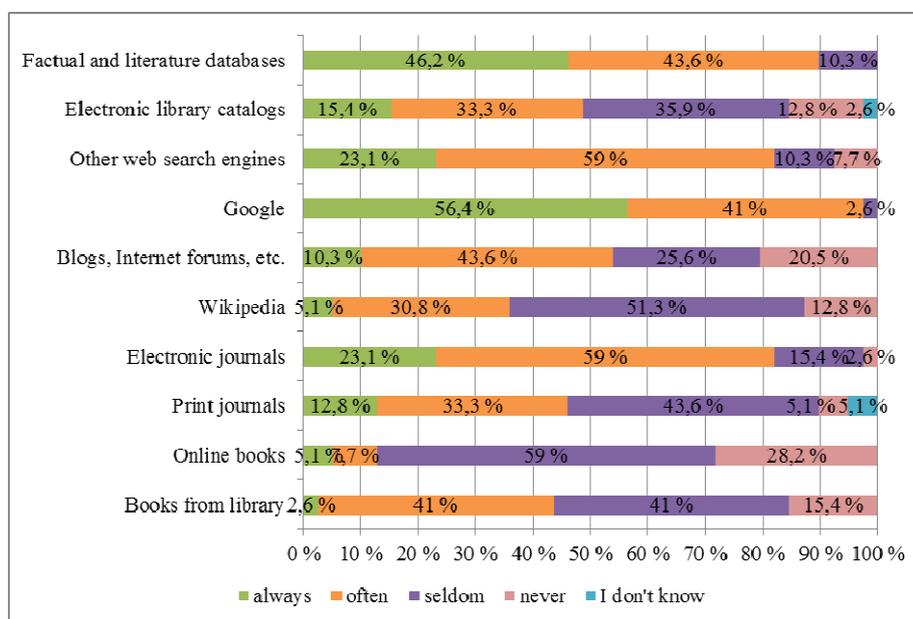


Fig. 3 Use of different information sources and tools by students from Chungbuk National University

In contrast, the Korean students use databases, blogs, internet forums, and other web search engines (than Google) more frequently (fig. 3). The latter can be explained in that Google does not have a monopoly in Republic of Korea unlike in Europe.

4 Discussion

Our study illustrates that there are big differences in the level of information literacy between master students from the University of Graz and CBNU. On average, students from the University of Graz are graded one level higher than CBNU students. Both Austrian and Korean students overly assess their information literacy skills in comparison with their actual skills resulted from the knowledge tests. This is particularly true for information activities which they are accustomed to, for instance, searching for information and defining the information need.

To some degree, these differences may be attributed to cultural and social issues. While Western societies are much more individualistic, Asian cultures usually have much higher levels of collectivism (Kim, Sohn & Choi, 2011). One of the most famous concepts for explaining intercultural differences was developed by Hofstede. In one of his more recent publications, Hofstede (2001) distinguishes between six dimensions which can be used to explain cultural differences: power distance, individualism/collectivism, masculinity/femininity, uncertainty avoidance, long-term/short-term orientation, and indulgence/restraint. Hofstede also provides a tool which delivers scores (0–100) for each of these six dimensions for a country.² For instance, Austria scores low (11 points) in the power distance dimension, while Republic of Korea has a much higher score (60 points). This means that Korea is a more hierarchical society.

Although this study does not directly measure the relationship between cultural aspects and information literacy, these cultural and social differences may influence the developments of education systems (Sung & Han, 2011). In the Korean education system, known as ‘cramming education’, it is common practice to absorb large volumes of given informational material. In this case, students need to memorize a lot of information to meet a highly required standard (ibid.). This means that the Korean education system does not focus on developing students’ own thinking and creativity but building a large knowledge deposit on students (ibid.). Meanwhile, the Austrian and many Western education systems focus on developing problem-solving capabilities of students (ibid.). As a result, this kind of norm may induce students to develop their information literacies. For instance, finding information possibly requires a certain degree of creativity of using and combining key words. If these key words are given, maybe, Korean students can find relevant information faster than Austrian students. However, this was not part of the information literacy test. Furthermore, this kind of cultural background explains why the most popular web search engines in Korea are Naver and Daum, but not Google. These web search engines suggest only the most relevant websites, while Google lists all websites to its users from which they have to choose the most relevant ones. If the information literacy test would include performance dimensions with regard to finding and organizing information, for example, how fast a student finds relevant information for given key words, maybe, Korean students would perform higher than Aus-

² <https://geert-hofstede.com/countries.html>

trian students. Although this speculation is one way to explain the results, this conjecture should be studied in more detail in the future.

Furthermore, the analysis concerning the use of information sources indicates that CBNU students are more IT literate than their counterparts from Graz. The use of search engines is part of everyday life in Korea. Moreover, CBNU students rely much more on “modern” media, like for instance blogs, internet forums, search engines, and databases. It follows that a high level of IT literacy does not go along automatically with a high level of information literacy. But again, this finding should be further studied along with the speculation mentioned above.

References

- American Library Association (1989): Presidential Committee on Information Literacy. Final Report. <http://www.ala.org/acrl/publications/whitepapers/presidential>
- American Library Association (2000): Information Literacy Competency Standards for Higher Education. <http://www.ala.org/acrl/sites/ala.org/acrl/files/content/standards/standards.pdf>
- Association of College and Research Libraries (2016): Framework for Information Literacy for Higher Education. <http://www.ala.org/acrl/standards/ilframework>
- Beutelspacher, Lisa (2014): Erfassung von Informationskompetenz mithilfe von Multiple-Choice Fragebogen. In: *Information. Wissenschaft & Praxis*, 65 (6), 341–352.
- Beutelspacher, Lisa, Maria Henkel, and Christian Schlögl (2015): Evaluating an Information Literacy Assessment Instrument. In: Pehar, Franjo et al. (Eds.): *Re:inventing Information Science in the Networked Society*, ISI 2015, 14th International Symposium on Information Science, Zadar, Croatia, May 19th–21st, 2015, Proceedings (pp. 482–491). Glückstadt: Hülsbusch.
- Eisenberg, M. and R. Berkowitz (1990): *Information problem solving: The Big Six skills approach to library & information skills instruction*. Norwood, N.J.: Ablex.
- Hofstede, Geert (2001): *Culture's Consequences: Comparing Values, Behaviours, Institutions and Organizations across Nations* (2nd ed.). Sage: Thousand Oaks.
- Kim, Yoojung, Dongyoung Sohn, and Sejung Marina Choi (2011): Cultural Differences in Motivations for Using Social Network Sites: A Comparative Study of American and Korean College Students. In: *Computers in Human Behavior*, 27 (1), 365–372.

- Klatt, Rüdiger, Konstantin Gavriilidis, Kirsten Kleinsimlinghaus, Maresa Feldmann, et al. (2001): Nutzung elektronischer wissenschaftlicher Information in der Hochschulausbildung. Barrieren und Potenziale der innovativen Mediennutzung im Lernalltag der Hochschulen. Dortmund. <http://ids.hof.uni-halle.de/documents/t267.pdf>
- Kuhlthau, C. (1991): Inside the search process: Information seeking from the user's perspective. In: *Journal of the American Society for Information Science* 42 (5), 361–371.
- Maurer, Angela (2016): *Informationskompetenz an der Universität Graz: Messung der Informationskompetenz von StudienanfängerInnen verschiedener Studienrichtungen*. Master Thesis, University of Graz.
- Maurer, Angela, Christian Schlögl, and Stefan Dreisiebner (2017): Comparing Information Literacy of student beginners among different branches of study. Paper submitted to Librellarium, <http://libellarium.org>, Paper is still under review.
- Rust, Kathrin (2016): *Comparing Information Literacy of Students from the University of Graz (Austria) and Chungbuk National University (Republic of Korea)*. Master Thesis, University of Graz.
- Sung, E. and Y. Han (2011): Implicit Thought of Korean and the US Students about Component Promoting or Hindering Creativity Development. In: *The Korean Journal of Educational Psychology*, 25 (4), 927–944.

Visualising Topics in Document Collections

An Analysis of the Interpretation Processes of Historians

Anastasia Christoforidis, Ben Heuwing, Thomas Mandl

University of Hildesheim, Germany

anastasiachristoforidou@gmail.com, {heuwing, Mandl}@uni-hildesheim.de

Abstract

This paper discusses two multivariate visualisations which provide insights into topic model distributions across sub-collections of a collection of historical textbooks in the context of a digital humanities project. Results of a qualitative user study with experts in historical research indicate that network graphs are more appropriate for revealing general connections among sub-collections, while small-multiples of heatmaps of topic correlations are more suitable for a finer grained analysis of the connections between specific topics. We analyse the user behaviour during analysis to identify general activities of the interpretation of topic models as well as activities of interpreting visual elements that are specific to each visualisation. As a result, we observed usability problems and show potential for improved visualisations in digital humanities applications.

Keywords: information visualisations; topic modelling; digital humanities; evaluation; user study

1 Introduction

In the humanities, access to large collections of digitised sources has created an increased interest into the use of tools to support the analysis of the contents of large amounts of documents. While in disciplines such as literature

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 37–49.

studies, analysis takes into account mainly stylistic characteristics of documents, historians largely depend on an analysis at a topical level to make sense of a collection and to retrieve documents of interest. Modelling a collection and providing interactive visualisations of its contents are important and interdependent steps that are necessary to offer support for the research interests of researchers in the humanities, many of whom do not have a background in information technology.

Topic modelling has become a common approach to text analysis in the digital humanities (Blei, 2012). Our contribution is based on a project which aims to provide support for the analysis of a collection of textbooks used in schools in Germany in the 19th and early 20th century. The collection is comprised of 3,803 textbooks with 799,260 pages. Metadata is available over a wide range of attributes, including subject, type of school, date, and place of publication.

Prior user research (Heuwing, Mandl & Womser-Hacker, 2016) led to the conclusion that, in addition to an open-ended explorative analysis to find patterns and trends of interest, the most important task consists of a comparative analysis of the contents of the collection across sub-collections, either according to a dimension of metadata of interest, over time, or both. This kind of analysis is focused on a specific area of interest, which may be represented by several similar topics of the topic model. To this end two prototypical visualisations were created, one displaying topics and sub-collections as nodes in a network graph, the second one making use of heatmaps to compare correlations of topics in sub-collections. These tools were comparatively evaluated with researchers working in history with the aim to examine ways of analysing provided information (including metadata, topics and further metrics denoting their interdependencies, cf. sect. 4) for the collection when applying the visualisations to solve realistic tasks. Thereby, we try to shed some light on the following, explorative research question: *To which degree do both visualisation tools, also with respect to the represented data, support data analysis processes?*

2 Topic modelling

Topic Modelling describes a set of algorithms which help to analyse a large collection of documents based on its latent thematic structure. The most frequently used technique LDA (Latent Dirichlet Allocation) assumes that every document in the collection is generated from a fixed number of topics, each document exhibiting a different proportion of each topic. Every topic is defined as a distribution over all words within the document collection as a fixed vocabulary, giving high weights to those words that tend to co-occur (Blei, 2012). These properties of topic modelling enable the annotation of documents, while the aggregated distributions of topics are expected to resemble the thematic structure of the document collection. Based on that, exploration and further information foraging tasks over the document collection can be supported (ibid.). The potential for the field of digital humanities lies in the verification and formation of theories about a document collection of interest. However, a topic model generated on a collection should not be understood as an objective representation of its contents, but rather as a “lens for viewing a corpus of documents” (DiMaggio, Nag & Blei, 2013: 582) which is specific to the focus of the researcher and her research interests.

3 Visualisation of topic models

Interactive information visualisations are necessary when analysing complex data sets in an effective and efficient manner. The represented data can trigger insights that are relevant for a domain. Specifically, for the interpretation of topic models visualisations can provide an overview over the thematic structure of a collection and help to reveal relations between topics and between topics and sub-collections.

Only few previous approaches to support the analysis of sub-collections based on topic models have been described in the literature. *DiTop* (Oelke et al., 2014) is a tool that makes use of glyph based and spatial techniques to support the comparative analyses of up to three different sub-collections. In this context, glyphs distinguish *discriminative topics*, which are distinctive and characteristic for one of these sub-collections from *common topics*,

which are characteristic for all documents. These representations of topics are positioned according to their representativeness for the sub-collections.

Visualisations using network techniques are very popular for displaying topic models in general. They not only reflect relations between different entities in an intuitive manner, but they are also easily extensible to additional visual dimensions and thus suitable for representing multivariate data (Gretarsson et al., 2012). An example that includes documents and sub-collections of a collection is *TopicNets*. Collections of documents and topics are represented as nodes of the network, while the distribution of a topic in a collection is conveyed through the connecting edges. The similarity between topics is represented through node positioning, so that clusters of documents that exhibit similar topics can be identified (ibid.).

4 Prototyping visualisation techniques to compare topics

The implementation of a visualisation design is constrained by the underlying data set and its attributes (e.g. the dimensionality and type, either qualitative or quantitative) and should be optimised to support the most important user goals. For the project, the primary goal is to assist historians in proving or rejecting hypotheses about the contents of a collection of text documents in terms of relations between different sub-collections that are defined by attributes of the metadata of the documents. Activities of analysis that are necessary to reach this goal may include the analysis of relations between a selected set of topics relevant to the analysis and the analysis of differences in the relations of these topics to sub-collections. Accordingly, the analysis of topic distributions in sub-collections and their changes over time is a major concept. Based on the available data and the user goals that have been identified, for the comparative evaluation study two alternative visualisation designs have been implemented: Network graphs and small-multiples.

Figure 1 shows an example of a network graph employed in the study: Nodes of different types either depict topics, labelled with the three most common terms, or sub-collections according to a selected attribute of the documents. The different node types are assigned to specific colours. The overall intensity of a topic in the collection and the number of documents in a

sub-collection are mapped onto node size. Edges are connecting only nodes of different types, i.e. topics and sub-collections. The width of edges encodes the average topic intensity in a sub-collection, i.e. the average of the values of a topic in all documents of a sub-collection. Considering the layout, the proposition of Gretarsson et al. (2012) is adopted to preserve both information of relations between topics as well as between topics and the sub-collections of the corpus. First, the position of the topic nodes is fixed on the two-dimensional plane according to the similarity of topics based on multi-dimensional scaling, and second, a common graph layout algorithm is applied to position the nodes representing sub-collections within the topic space. According to this template, static graphs representing four different time intervals within the span of twenty years have been generated to analyse changes over time.

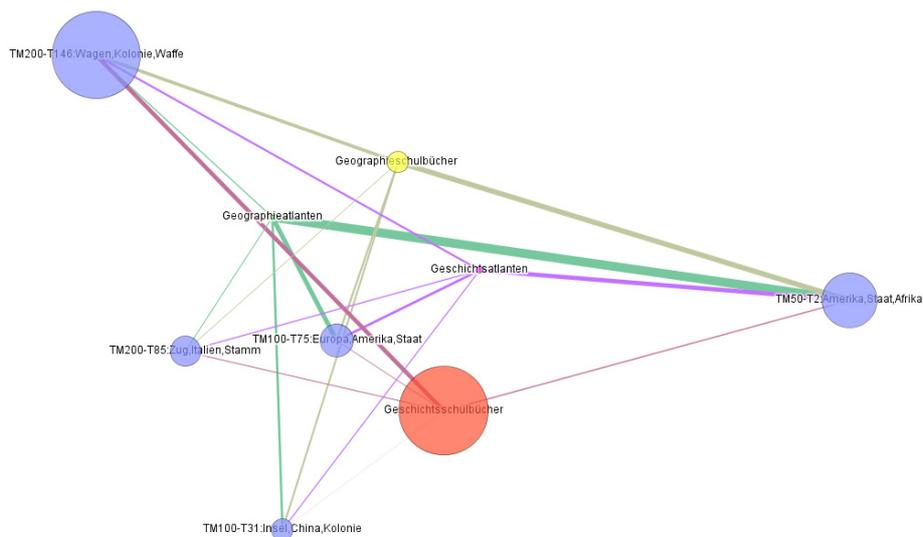


Fig. 1 Network graph consisting of topics and document sub-collections: blue nodes depict topics, other colours represent sub-collections (red for history textbooks, green for geographical atlases, purple for historical atlases, and yellow for geography textbooks)

The second design developed (cf. fig. 2) uses small-multiples to give a different perspective on the same data objects, by shifting the focus from relations between topics and document sub-collections towards relations between topics. Small-multiples are made up from small diagrams of the same

type and scale arranged in a grid according to different categories (Theus & Urbanek, 2008).

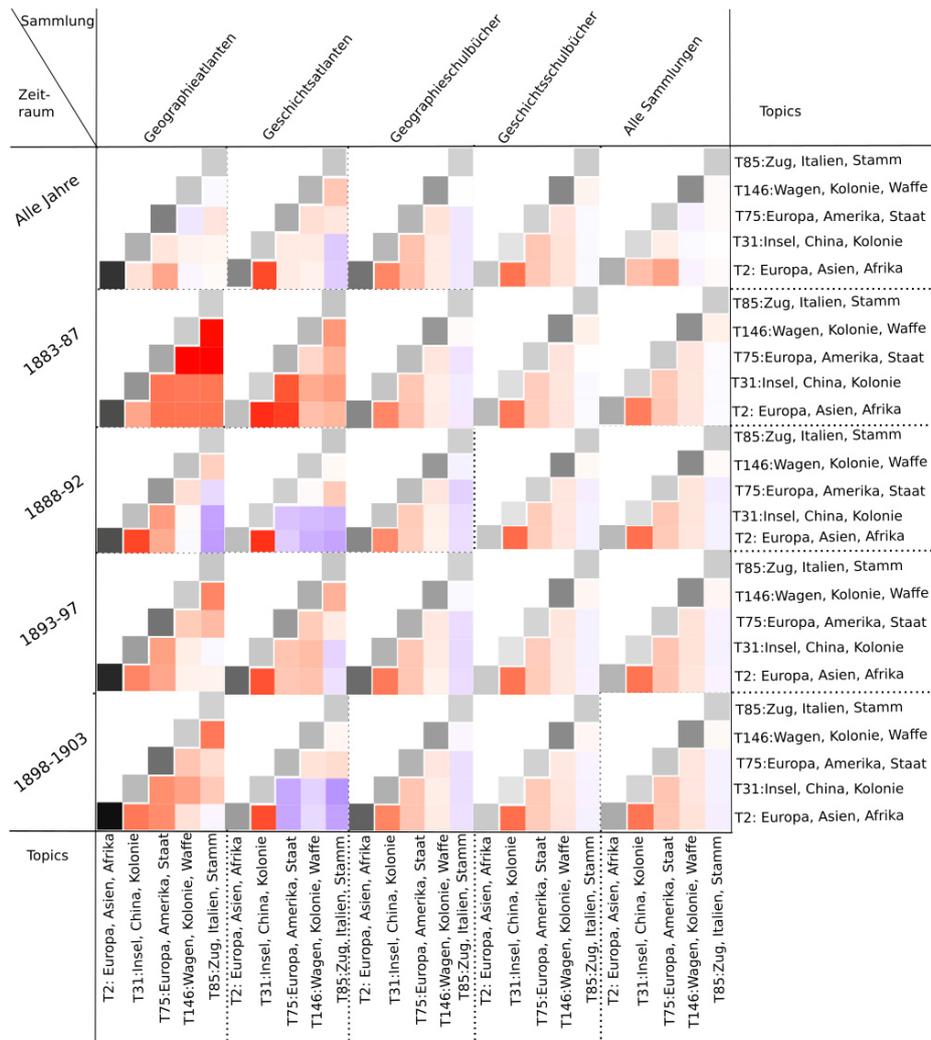


Fig. 2 Small-multiples consisting of heatmaps of topic correlations in sub-collections (columns: geographical atlases, historical atlases, geography textbooks, history textbooks, all collections; first row all years, followed by time intervals of five years)

In our visualisation, each consists of a matrix displaying the correlation of topics in the sub-collections. Pairwise correlation scores between topics were computed based on topic distributions over the documents for every interval in the time span of twenty years and for each sub-collection, as well as for the whole time span and over all sub-collections. Heatmaps were generated based on the resulting correlation matrices. As these are symmetrical, the upper half including the diagonal was left blank. Values from the respective correlation scores were mapped on an ordered, divergent colour map (ranging from blue for negative correlations over white for no correlations to red for positive correlations) to best represent the degree of correlation between two topics. The resulting heatmaps were then arranged in a grid according to their associated categories (time interval and document sub-collection). To maintain the information about the average topic intensity in a sub-collection, the rejected diagonal in the prior step was used by mapping this information onto an additional ordered, sequential colour map (using greyscales ranging from bright for very low topic intensity to dark for high topic intensity).

5 Methods

Task-based evaluation of visualisations tools can be used to study the fit between users' representation of the concepts and the representation chosen for the visualisations. Especially, when including both high-level tasks defined at a level of the users' goals in the domain and low-level tasks targeting the interaction with features of the visualisations, insights into the users' experience with the visualisations can be effectively gained (Faisal et al., 2007).

For the comparative evaluation of the two prototyped visualisation tools in the context of the above mentioned project, an empirical, task-based user study is conducted to capture user's experience from the target domain while interacting with the tools and the information conveyed by topics. Tasks are developed with analogies for both visualisations to be comparable. In the study, all users are presented with two low-level tasks for each tool to understand how they perceive visual elements and whether these are mapped on the intended data and attributes. General examples include "Which three topics are most likely commonly addressed over all sub-collections and which one distinguishes most from the rest?" and "Compare average topic intensi-

ties in the sub-collections history textbooks and geography textbooks.” In addition, two high-level tasks for each tool address the characteristic of tasks in real conditions, e.g. being open-ended and focusing on the interpretation in the context of research questions, for example “How can the relations between topics during the time span of twenty years (1883–1903) be described? Based on that, what can be concluded on the development of the co-occurrence of topics addressing colonies and emigration over all sub-collections?” and “How can the relation of topic T85 to others at the course of the whole time span (1883–1903) be described between the sub-collections geography textbooks and history textbooks? Based on that, what can you conclude on the development of relations between topics addressing emigration and colonies within history textbooks compared to geography textbooks?” Task order was randomised with regard to the tools, with low-level tasks for each tool being presented before the high-level tasks.

Observations and the participants’ explanations of their behaviour (“thinking aloud”-protocol) gave insights into the reasoning processes of participants during task completion. Results were selectively transcribed along with notes based on observations.

Five participants were selected in a high quality sample of expert users from the potential target groups of researchers in history. Each session took approximately an hour. At the beginning of each session, the rationale behind the study, topic modelling and the visualisations that were used were introduced. Every tool was provided as a static visualisation on the screen.

6 Results

Qualitative results were classified by making use of a systematic process of content analysis, thereby deriving a category system (Mayring, 2002). The following sections first present general activities of topic model interpretation, followed by the specific steps of interpreting the visual elements of the two visualisation tools together with common problems that occurred.

6.1 General activities when interpreting topic models

For the completion of tasks, participants applied different activities to understand the underlying models, i.e. topics in relation to other topics, their connections to sub-collections, and relating them to their own existing contextual knowledge. In the study, three participants derived new knowledge about the domain from the model based on these activities combined with prior conducted activities regarding the visual elements of the respective visualisations. For example, they draw a conclusion on the subjects addressed in a particular sub-collection, like one participant concluded on the content of history textbooks: “History textbooks differentiate between antique emigration and modern emigration along colonies.”

Understanding relations between topics within sub-collections (2): The analysis of topics with similar or different distributions to each other was an important aspect for the interpretation of topics and their relations within sub-collections. A recognised similarity or dissimilarity between two topics within a sub-collection was explained by interpreting the content of topics using topic terms and then relating them to the features of the sub-collection. Hence, a low similarity of a topic to others within the sub-collection containing geography textbooks is explained by the fact that these are less likely to deal with socio-cultural topics, as one participant mentioned: “The topic [with the terms wagon, colony, weapon] does not play an important role within geography textbooks. [...]. I assume that these address more topics similar to contents of geography than migrations, tribes and ethnicities.” In one case, a participant recognised a high correlation between the topics with the terms *wagon*, *colony*, *weapon* and *migration*, *italy*, *tribe* in the sub-collection containing history textbooks, and concluded that history textbooks predominantly address the subjects of migration during the time of the Roman Empire rather than in the context of German colonies.

Contextual assignment of a topic to a sub-collection (2): Instead of taking only the represented data into account to relate a topic to a sub-collection, some participants referred to their own contextual knowledge. Consequently, a participant explains her own association of a topic to geography textbooks: “Regardless of the visualisation, I would connect [the topic containing the terms] europe, america, state more to geography textbooks than migration, italy, tribe. [...] My association would be to connect geography textbooks in terms of the content more to states, than migration, italy, tribe. Tribe is something cultural and can’t be classified into a geographical context.”

Referring to topics: Topics were mostly referred to by the identifiers used in the visualisation. One participant explicitly named a topic antique topic, based on the interpretation of its terms, with the three most common *migration, italy, tribe*, as she explains her reasoning: “This [topic] addresses migration rather than emigration, the Italy campaign of Goths and Vandals [...]. This is thousand years prior to these [topics].” In one case, the participant also named a topic *main geography topic*, based on its high intensity in the sub-collection of geography textbooks, thus transferring characteristics of the sub-collection to the topic.

6.2 Comparison of visualisations

The following categories describe activities conducted by the participants when interpreting the visual elements of the visualisations. Problems occurred especially at the beginning of tasks. During the test session, the participants’ understanding of the visualisations and their confidence when interpreting them increased.

6.2.1 Network graphs

Process of gaining insight and generating information: Mostly, the visual elements, e.g. nodes or edges, were identified first. Then, further relevant attributes of these elements were interpreted, most importantly average topic intensity in a sub-collection conveyed through edge width, and topic similarity through position and proximities of topic nodes. Next, participants compared these attributes across time intervals. In one case, edge widths of all connections of a sub-collection node to all topic nodes were aggregated to describe the overall topic intensity in a sub-collection. These steps lead to the generation of information, like relations between topics or between sub-collections and topics (2), similarities or differences between attributes of topics or the average topic intensity (3), and their development over the time intervals (3).

Problems: Regarding the use of visual elements, problems were caused by the encoding of colour (e.g. assigning colour use to node types) (2), positions (e.g. meaning and arrangement of nodes in the plane was not always clear) (2) or the distinction between the encoding of topic similarity and average topic intensity (1). In the high level tasks, participants often created mental models that did not always match the intended meaning of visual elements.

For example, centrality of topic nodes was interpreted as their overall relevance (1), while closeness was used to describe the relations between topics and sub-collections instead of using edge widths (2). Position was also applied to describe changes over time and the development of relations (2). Edges and their widths were not only used to describe the relation of topics and sub-collections, but also the relation between topics. In one case, a topic node with comparatively narrow edges to all of the sub-collection nodes was interpreted as not being relevant in the context of any other topic.

6.2.2 *Small-multiples*

Process of gaining insight and generating information: The initiation of an analysis usually consisted of identifying the colour map, either for topic correlations or average topic intensity (2). Next, participants tried to identify relevant categories. They either considered all categories in each time interval and sub-collection or used only the overview over the complete time span or all sub-collections. Relevant correlations between topics within specific categories were first identified in general (4). Similar to the use of network graphs, results of average topic intensities were sometimes aggregated, e.g. for a sub-collection (2). Correlations of several identified categories were compared next. A sub-activity here was to understand the impact of a correlation between two topics in a particular sub-collection on those in all sub-collections. Comparing different categories that way also resulted in finding interesting patterns (similarities or differences) regarding relations between topics and categories (either sub-collections or time interval or both) or correlations between topics (5). Investigating these patterns particularly resulted in recognising changes over time (2).

Problems: Commonly faced problems included matching the right topic pairs to each cell in the heatmap (3), recognising differences in saturation and luminance of the colour maps for precisely discriminating correlation scores (4) and creating an understanding for the generation of heatmaps by the underlying computation of the correlation scores (2).

6.3 Summary

Processes of analysis appear to be influenced by the form of visualisations. When compared to network graphs, small-multiples enable finer grained approaches for analysis and mainly support the generation of domain related

knowledge through the representation of correlations between a pair of topics in a particular sub-collection. Network graphs, on the other hand, tend to support more insights and provide a comprehensive overview of the relations of topics and sub-collections. Participants seem to readily interpret every visual feature to derive conclusions about relationships in the data, especially if their findings are congruent with existing background knowledge. But using complex graphs that include many visual dimensions also entails risks because elements of the visual syntax are more likely to be misinterpreted.

7 Conclusion

The participation of representatives of the very specialised domain of historical textbook researchers together with the qualitative analysis of results provide insights into the mental processes applied during the analysis which are highly relevant for the project. As the results show, historians reveal many approaches of analysing information about topic models, where topic terms play a major role. Activities in this context are also mainly responsible for deriving general conclusions on the content of particular sub-collections compared to others and thus creating domain related knowledge.

The analysis of the activities for the interpretation of the models and visualisations highlight opportunities for improving our visualisations of topic models. For example, the strategies to compare distributions and correlations of topics within sub-collections to the respective values of the complete collection indicate the need for comparative metrics (e.g. higher/lower than general), while the important role of external, contextual knowledge indicates the need to annotate elements with the users' own terminology. Qualitative findings reveal that network graphs are less difficult to understand but may also lead to mis- and over-interpretations of the data provided by topic models. This finding is probably also relevant for other areas where visualisations of topic models are used as a basis for further interpretations.

Acknowledgment

We would like to thank all the participants of the study for their support and input. This research was partly funded by a SAW grant from the Leibniz Association.

References

- Blei, D. M. (2012): Probabilistic Topic Models. In: *Communications of the ACM*, 55 (4). <http://www.aaii.org/ocs/index.php/ICWSM/ICWSM12/paper/download/4645</5021>
- DiMaggio, P., M. Nag, and D. M. Blei (2013): Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. In: *Poetics*, 41 (6), 570–606. <http://linkinghub.elsevier.com/retrieve/pii/S0304422X13000661>
- Faisal, S., P. Cairns, and A. Blandford (2007): Challenges of Evaluating the Information Visualisation Experience. In: *BCS-HCI '07, Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI ... but not as we know it – Volume 2* (pp. 167–170). University of Lancaster: BCS Learning & Development Ltd. Swindon. <http://dl.acm.org/citation.cfm?id=1531407.1531451>
- Gretarsson, B., J. O'Donovan, S. Bostandjiev, T. Höllerer, A. Asuncion, D. Newman et al. (2012): TopicNets: Visual Analysis of Large Text Corpora with Topic Modeling. In: *ACM Trans. Intell. Syst. Technol.*, 3 (2). <http://doi.acm.org/10.1145/2089094.2089099>
- Heuwing, B., T. Mandl, and C. Womser-Hacker (2016): Combining contextual interviews and participative design to define requirements for text analysis of historical media. In: *ISIC: The Information Behaviour Conference*. Zadar
- Mayring, P. (2002): *Einführung in die qualitative Sozialforschung: eine Anleitung zu qualitativem Denken* (5th ed.). Weinheim and Basel: Beltz.
- Oelke, D., H. Strobel, C. Rohrdanz, I. Gurevych, I., and O. Deussen (2014): Comparative Exploration of Document Collections: A Visual Analytics Approach. In: *Computer Graphics Forum*, 33 (3), 201–210. <http://dx.doi.org/10.1111/cgf.12376>
- Theus, M. and S. Urbanek (2008): *Interactive Graphics for Data Analysis: Principles and Examples*. Boca Raton, FL: Chapman & Hall/CRC.
- Zugal, S., J. Pinggera, H. Reijers, M. Reichert, and B. Weber (2012): Making the Case for Measuring Mental Effort. In: Chaudron et al. (Eds.): *EESSMod '12, Proceedings of the Second Edition of the International Workshop on Experiences and Empirical Studies in Software Modelling* (pp. 37–43). New York: ACM. <http://doi.acm.org/10.1145/2424563.2424571>

An Eye-Tracking Study on Differences in Information Transfer by Infographics

*Barbara Ströhl, Kilian Ganz, Stephanie Richter,
Kilian Zieglmeier, Rainer Hammwöhner †*

Chair for Information Science, Regensburg University, Germany
barbara.stroehl@ur.de,
{kilian.ganz, stephanie.richter, kilian.zieglmeier}@stud.uni-regensburg.de

Abstract

Information graphics are commonly used to display information. Nevertheless, the retention of information can differ depending on the presentation of the content.

A topic that is currently present to all of us in the media is the refugee influx to Europe. As it caused a lot of chaos, confusion and anxiety, the transfer of information played and still plays a crucial role, which is why we chose two different graphics visualizing facts and information about refugees.

Our aims were to get insights into the readers' information behavior dealing with information graphics and to find differences in information transfer.

Therefore, we conducted eye-tracking experiments and analyzed the fixation time and the fixation count on both textual and non-textual elements of the infographics. After reading, the retention of information was tested using free text questions and summed up in a score that was evaluated.

Results showed that the subjects had spent most of their time on textual elements for each infographic. The viewing behavior did not differ significantly between the two graphics. Despite this, we found significant differences in information transfer. This might be because one infographic had fewer sub-topics. Each of these sub-topics was backed up by the repetition of several textual and non-textual elements as well as additional details, which broadened the context.

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 50–61.

Keywords: eye-tracking; information graphic; textual elements; pictorial elements

1 Introduction

Last year, the German government had to deal with more than 650,000 asylum applications in the period from January to September. This represents an increase of about 134% in comparison to the same period for the previous year.¹ People from the Near East, Africa or other crisis areas had fled from civil wars and violence. The arrival of the refugees ignited political debates, xenophobia and even riots in some countries. In contrast, many people are willing to help the fugitives and try to understand their problems.

To do so, it is essential that information about the origins, the backgrounds and the motives of the refugees is provided. There are many different ways to transport information, e.g. plain text articles, images or videos. In this study, we focused on a special combination of text and visualization: infographics. The following paper deals with two questions. How do readers interact with these two components of infographics and does the way the information is displayed make a difference regarding the transfer of knowledge?

2 Related work

The background of this current issue is how to use the transfer of information and knowledge in an appropriate way. Users should get an easy and quick overview of purified and well-arranged information in order to fulfil their information need. Growing data sets produced by statisticians or collected by supercomputers provide the need of a suitable presentation form (Schumann, 2000). For many years, information graphics have been in use now to display

¹ Bundesamt für Migration und Flüchtlinge, “Aktuelle Zahlen zu Asyl” 09/2016, http://www.bamf.de/SharedDocs/Anlagen/DE/Downloads/Infothek/Statistik/Asyl/aktuelle-zahlen-zu-asyl-september-2016.pdf?__blob=publication <11.11.2016>

data and facts in a simple and comprehensible way (Bouchon, 2007). Bouchon defines an infographic as a combination of graphical elements, e.g. photos, drawings or pictograms, which draw attention and convey information that is perceivable at first glance, and typographic components like letters, digits, or mathematical symbols, which point out connections, functions and chronological sequences. For her, it is only through this combination that infographics convey additional information (ibid.). Especially daily newspapers frequently apply this way of visualizing information. The increasing popularity of infographics supports their claim to represent a good way to transfer and report information.

Information graphics are the object of scientific investigations often. Especially their possibilities to transport and display information and their usefulness in the communication of the digital age have been analyzed for numerous times.

A study by Holmqvist and Wartenberg (2005) showed that the presentation of information had an impact on reading behavior. They compared a graphic with integrated text and images and a graphic with separated text and pictures. Results showed that readers focused more on the images when both elements were separate. When the pictures were integrated in the text, readers looked at the images and read the text equally (ibid.).

Based on these results Holsanova et al. (2009) recorded eye-movements to evaluate the interaction of readers with given information graphics. They additionally used different design laws such as spatial contiguity or arrows pointing towards related elements to evaluate the impact on the understanding of the context and information transfer. The evaluation of the eye-tracking data showed that readers tend to jump from the headline to graphic elements directly when text and graphics are separate. In comparison to this, readers who had to read the text with integrated graphics, both elements, pictures and text blocks, were read together (ibid.).

Another study by Dagmar Gehl (2012) compared two magazine articles, which were either original or manipulated a distinct way, in order to find influences on the knowledge transfer. The original version of each article consisted in textual and graphical elements. In the modified version, the important pieces of information from pictorial elements were converted into text and the images were removed. For assessment, every subject had to fill in a questionnaire and a concept map. The results proved that in both parts of the test, the participants that had to read the unmodified article with graphics achieved better results. The eye-tracking data supported this.

Further analysis proved that graphics catch the reader's eye. A redundant information presentation and a strong interaction with the content influence the information transfer in a positive way (ibid.).

However, information graphics do not automatically transfer knowledge. They can only be helpful, if the user assess the information as relevant for his information need (Burmester & Wenzel, 2013).

Information graphics are always created in a user-centric design. This creates a gap between user and creator: The user wants to fulfil his information needs relating to a current subject, while the designer tries to convey specific information (ibid.).

The related work shows that recipients remember information better when textual elements are enriched and even integrated by graphical elements. However, it is important to note that each person interprets things differently and that information has to be assessed as relevant by the reader to be remembered.

3 Experiment design

The information behavior was assessed in an eye-tracking experiment on subjects reading infographics. The transfer of information was measured by a score based on questions about information covered by each graphic.

3.1 Selection of stimuli

There are many information graphics that transfer details about the numbers of refugees, their migration paths, and numerous other facts about fugitives. The stimuli to be analyzed were selected by the following criteria:

As we focused on differences in information processing, the graphics had to be comparable, but not too similar. Furthermore, the graphics needed to contain textual elements enriched with images or pictograms. The textual elements should not be longer than a few lines, and images or other non-textual elements, such as charts, diagrams or maps had to be memorable but not too dominant. Moreover, a balanced mixture of textual and non-textual elements was considered ideal for our research purposes and the overall quantity of textual elements had to be similar for both graphics.

In order to assess the information transfer, a common questionnaire had to be used, so each graphic had to cover all the information needed to answer it. We chose two graphics from the web presence of a German governmental institution², which was considered a reliable source concerning the content.

The two selected infographics differ mostly in their appearance: The first graphic “*Flucht in Zahlen*” (fig. 1) can be roughly translated as “Refugee Facts in Numbers”, and is referred to as ‘NUMBERS’ in this text. This graphic is designed in light colors like grey and white. Textual elements are arranged around rather big images, marking the center of the graphic. In addition, big numbers symbolize important values. This graphic contained a total amount of 498 words.



Fig. 1 ‘Flucht in Zahlen’ (‘Refugee Facts in Numbers’, referred to as ‘NUMBERS’)³

The second information graphic “*Der Weg über das Wasser*” (cf. fig. 2) means “The Route over the Water” and is referred to as ‘ROUTE’. This graphic is kept in darker colors such as black, blue and red. Textual elements are scattered all over the graphic. Instead of concise images, pictograms are used to reveal information about related textual elements. ‘ROUTE’ has a count of 408 words.

2 Bundeszentrale für politische Bildung (Federal Agency for Civic Education)
 3 <http://www.bpb.de/shop/zeitschriften/213674/bpbmagazin-2-2015> <15.11.2016>

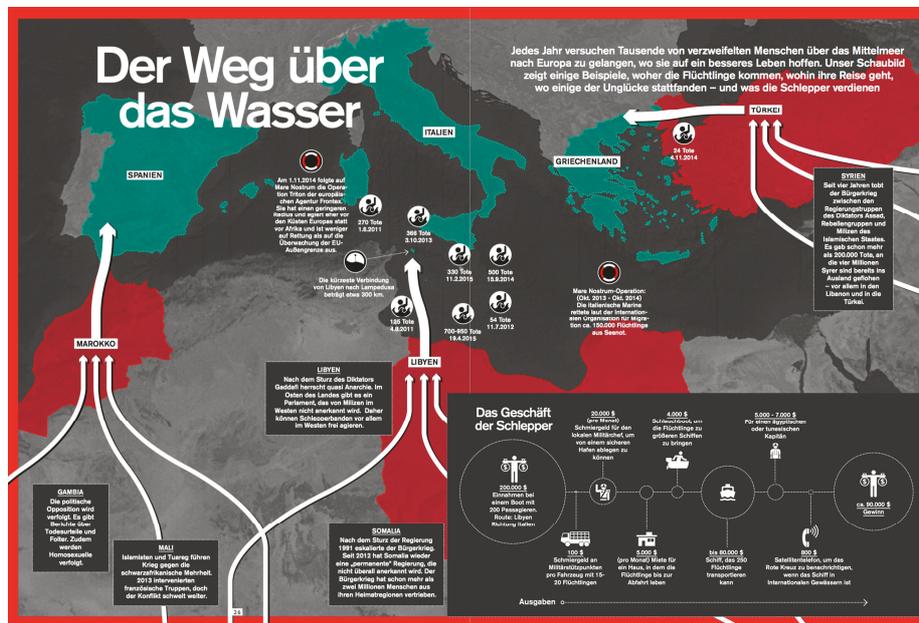


Fig. 2
 ‘Der Weg über das Wasser’ (‘The Route over the Water’, referred to as ‘ROUTE’)⁴

3.2 Assessment of information transfer

A questionnaire was designed to assess the amount of information transfer. It contains questions about information that both ‘NUMBERS’ and ‘ROUTE’ cover. Some pieces of information are present in textual elements of each graphic, whereas others are displayed evidently in the center of ‘NUMBERS’ but wrapped up in text in ‘ROUTE’. According to Bloom’s learning target taxonomy, we focused on the category ‘knowledge’, i.e. remembered facts (Bloom & Engelhart, 1976). The subjects were asked three questions to be answered in free text.

The first question referred to the title of the graphic. The information required to answer questions two (death count of refugees) and three (earnings of smugglers) was depicted in the text and in corresponding pictograms. We additionally asked the subjects demographic questions like age and gender. The collection of these data was carried out via Google Docs.

4 <http://www.bpb.de/shop/zeitschriften/fluter/208588/flucht?blickinsbuch> (pp. 26–27) <15.11.2016>

3.3 Experimental procedure

Each subject was randomly shown either ‘NUMBERS’ or ‘ROUTE’. The experimental setup took place in the eye-tracking laboratory at the chair of information science. Data was recorded using a SMI 250 Hz remote eye-tracking device. Every subject was placed at a distance of about 60 to 70 cm in front of the screen. After instructions and calibration the graphic was displayed. The participant had no time limit while reading the information graphic. The last step was to answer questions about the presented content.

After two pre-tests for adjusting the experimental set-up we tested 30 subjects. 15 subjects were shown the graphic ‘NUMBERS’ or ‘ROUTE’, respectively. Their age ranged between 15 and 62 years with an average age of 27.3 (SD = 11.9) years. 18 of the subjects were male and 12 were female.

All subjects were familiar with the topic. Each of them had either graduated from high school or had a university degree.

3.4 Data preparation

We focused on the reading behavior on text snippets and the viewing behavior on images. Within each infographic, we separated the textual and the non-textual elements into two different groups of areas of interest (AOIs) in order to compare these elements.

We determined different parameters like the fixation time in seconds and the fixation count, i.e. the number of fixations, for both groups of AOIs (textual and non-textual elements) as well as the first and the second fixation on the stimuli. The fixation time tells us how long the reader’s eye has remained on a specific area and is the most frequently reported parameter in eye-tracking research (Holmqvist et al., 2011). The fixation time is associated with cognitive processing according to the eye-mind hypothesis (Just & Carpenter, 1980). That means the longer an area is fixated the deeper is the information processing. The fixation count tells us the number of individual fixations per AOI. In combination with the fixation time, this provides additional information. For example, whether an area is fixated frequently but for a short time or consists of few fixations with a long duration (Holmqvist et al., 2011). Longer fixation time and a higher fixation count refer to a deeper understanding of the text (ibid.). Concerning the viewing behavior on pictorial elements, higher fixation time and count indicate information that is considered relevant by the viewer (Loftus & Mackworth, 1978).

Furthermore, we investigated starting points. The first and second fixation on a stimulus tells us which elements are most salient. The first entry point is unconscious and not influenced by characteristic properties of the stimulus. The second fixation, however, is executed actively and reflects the reader's processing (Holmqvist et al., 2011).

To evaluate the retention of information, the results of the questionnaire were summed up in a score. Each correct answer was credited with two points. When asked for a number, the subject was credited one point when his answer was within a range of plus or minus five percent of the correct number. No answer or wrong answers were credited with zero points. The answers were summed up in a normalized score (number of points obtained divided by maximum points) between 0 and 1.

4 Data analysis

Firstly, we wanted to check if information transfer is different between the two graphics. The created score ranges between 0 and 1 and the level of measurement is metric. A Shapiro-Wilk test showed a non-Gaussian distribution of the score for both groups (p -values $< .05$). In order to compare the scores, we used the non-parametric Mann-Whitney U test. We found a significant difference between the two groups ($Z = -3.1854$; p -value $< .01$). The mean score for the graphic 'NUMBERS' was 0.17 (SD = 0.19), 'ROUTE' had an average value of 0.42 (SD = 0.20). Interestingly, the median of the infographic 'NUMBERS' was at zero which means that half of the participants did not score a single point in the questionnaire. The second graphic 'ROUTE' on the other hand showed a score of 0.4. In table 1 the dispersion of the values is shown.

Table 1: Score summary for each graphic

graphic	min	25 th percentile	median	mean	75 th percentile	max
'NUMBERS'	0.0	0.0	0.0	0.17	0.33	0.5
'ROUTE'	0.0	0.4	0.4	0.42	0.42	0.8

The mean score of the two infographics differed by 0.25. Keeping in mind the overall range of just 1, this also supports the hypothesis of different ways

of information transfer in the two graphics. Accordingly, the 75th percentile and the maximum were much higher in the infographic 'ROUTE'.

To sum it up, the participants remembered more information from the graphic 'ROUTE'.

To understand why the participants had a higher rate of information transfer from one of the graphics, we scrutinized their way of reception by determining several eye-tracking parameters on certain areas of the infographics.

In general, textual elements were fixated longer than non-textual elements in both infographics. For 'NUMBERS' the average fixation time of textual elements was three times higher than the one of non-textual elements. On average, the textual elements were fixated for 99.82 seconds (SD = 69.31 sec), the pictorial elements were fixated 33.39 seconds (SD = 14.65 sec). The same phenomenon was observed in the graphic 'ROUTE' with average values of 92.56 seconds for textual elements (SD = 35.84 sec) and 26.70 seconds for pictures (SD = 9.59 sec). The fixation count corroborated these results. The number of fixations on textual elements (M = 182.3, SD = 90.87) in the graphic 'NUMBERS' was higher as the count on the images (M = 106.7, SD = 43.75). The mean fixation count on textual elements in the graphic 'ROUTE' was 182.4 (SD = 45.34) and on non-textual elements 77.5 (SD = 31.74). A subject spent on average 2.22 minutes (SD = 71.9 sec) on 'NUMBERS' and 2.07 minutes (SD = 48.0 sec) on 'ROUTE'. Due to different sizes of the AOIs, the fixation time of each was normalized by division through the total fixation time to grant comparability.

We checked for differences between the two infographics regarding the whole of textual and non-textual AOIs, respectively. Non-parametric Mann-Whitney U tests were used as there was no Gaussian distribution either as shown in a Shapiro-Wilk test (p -values < .001).

First, we compared the fixation time in percent on textual elements between the two infographics. There was no significant difference between both groups ($Z = -0.477$; p -value > 0.05) with the medians of 78.44 percent on 'NUMBERS' vs. 79.65 percent on 'ROUTE' as shown in table 2. So the participants spent about the same amount of time on textual elements in the graphics 'ROUTE' and 'NUMBERS'.

Additionally, we had a look at the normalized fixation time on pictorial elements. As well as the normalized fixation time on textual elements, there was no significant difference ($Z = 0.477$; p -value > 0.05) between the two graphics ('NUMBERS' median: 21.56; 'ROUTE' median: 20.35).

Table 2: Summary for normalized fixation time on AOIs

graphic	min	25 th percentile	median	mean	75 th percentile	max
textual elements						
“NUMBERS”	13.99	64.88	78.44	68.36	83.06	84.56
“ROUTE”	46.59	74.76	79.65	76.64	82.54	89.42
non-textual elements						
“NUMBERS”	15.44	16.94	21.56	31.64	35.12	86.01
“ROUTE”	10.58	17.46	20.35	23.36	25.24	53.41

We also analyzed the reading intensity by dividing the fixation time on textual AOIs (in seconds) for each participant through the number of words in the respective graphic. Student’s t-test was used in this case as the data were distributed Gaussian (Shapiro-Wilk test, $p > 0.05$) with homogenous variances (Bartlett test, $p > 0.05$). The test revealed no significant differences but the average reading intensity was slightly higher in the graphic ‘ROUTE’ ($M = 0.24$; $SD = 0.11$) than in ‘NUMBERS’ ($M = 0.2$; $SD = 0.13$).

Table 3: Summary for reading depth

Graphic	min	25 th percentile	median	mean	75 th percentile	max
“NUMBERS”	0.01	0.13	0.18	0.2	0.26	0.55
“ROUTE”	0.06	0.15	0.21	0.24	0.33	0.49

The last parameters we analyzed were the subjects’ first and second looks at the graphics. The analysis revealed that 90 percent of the first fixations were located in the center of the infographic. These results go along well with Tatler (2007), who stated that people tend to fixate the center of stimuli presented on a computer screen. We furthermore examined the subjects’ second fixations. About 80 percent of the participants moved their eyes from the center to the heading of the graphic. In this study, both of the headings were placed in the top left corner and had about the same size and the same color.

5 Discussion

To sum it up, textual elements received more attention than non-textual elements in both infographics. Images might be perceived at first glance whereas more time is needed to read a text paragraph.

Although no significant differences in information behavior could be stated, we observed a significant difference in information transfer between the two graphics. The subjects remembered more information from the graphic 'ROUTE'. This might be because fewer sub-topics are covered by this infographic. Each of these sub-topics is backed by either the repetition of textual and non-textual elements or additional details, which broaden the context. Another point is that 'ROUTE' mostly represents information with the help of pictograms, which are used repeatedly. The infographic 'NUMBERS', on the other hand, uses many different photos of real objects to illustrate its great variety of sub-topics. These differences reasonably explain the fact that the subjects could remember more information, but this study provides no data to either prove or disprove these hypotheses. What this study showed, however, is that the sole investigation of the viewing behavior is not sufficient to explain differences in information transfer. Therefore, confounding parameters need to be reduced, e.g. by using just one infographic that is displayed in different ways. Furthermore, plain text might represent an appropriate way to determine a sort of baseline. Besides that, the group of subjects needs to be increased either in numbers or, at the expense of external validity, in homogeneity. To sum it up, more variables than just the viewing behavior need to be considered to investigate information transfer in infographics.

References

- Bloom, B. S. and M. D. Engelhart (1976): *Taxonomie von Lernzielen im kognitiven Bereich*. Beltz: Weinheim.
- Bouchon, C. (2007): *Infografiken. Einsatz, Gestaltung und Informationsvermittlung*. Boizenburg: Hülsbusch.

- Burmester, M. and A. Wenzel (2013): Ansätze zur Evaluation interaktiver Infografiken. In: W. Weber et al. (Eds.): *Interaktive Infografiken* (pp. 85–104). Berlin: Springer.
- Gehl, M. D. (2012): Concept Mapping und Eyetracking: Eine Methodenkombination zur Diagnose medial initiiertes Wissenszuwächse. In: H.-J. Bucher and P. Schuhmacher (Eds.): *Interaktionale Rezeptionsforschung* (pp. 135–155). Wiesbaden: Springer Fachmedien.
- Holmqvist, K., and C. Wartenberg (2005): The role of local design factors for newspaper reading behaviour – an eye-tracking perspective. In: *Lund University Cognitive Studies*, 127, 1–21.
- Holmqvist, K. et al. (2011): *Eye Tracking. A Comprehensive Guide to Methods and Measures*. New York, NY: Oxford University Press.
- Holsanova, J., N. Holmberg, and K. Holmqvist (2009): Reading information graphics: The role of spatial contiguity and dual attentional guidance. In: *Applied Cognitive Psychology*, 23 (9), 1215–1226.
- Just, M. A. and P. A. Carpenter (1980): A theory of reading: From eye fixations to comprehension. In: *Psychological Review*, 87 (4), 329–354.
- Loftus, G. R. and N. H. Mackworth (1978): Cognitive Determinants of Fixation Location During Picture Viewing. In: *Journal of Experimental Psychology: Human Perception and Performance*, 4 (4), 565–572.
- Schumann, H., and W. Müller (2000): *Visualisierung: Grundlagen und allgemeine Methoden*. Berlin: Springer.
- Tatler, B. W. (2007): The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. In: *Journal of Vision*, 7 (14): 4, 1–17.

Are Ads on Google Search Engine Results Pages Labeled Clearly Enough?

The Influence of Knowledge on Search Ads on Users' Selection Behaviour

Dirk Lewandowski, Sebastian Sünkler, Friederike Kerkmann

Hamburg University of Applied Sciences, Germany
{dirk.lewandowski, sebastian.suenkler, friederike.kerkmann}@haw-hamburg.de

Abstract

In an online experiment using a representative sample of the German online population ($n = 1,000$), we compare users' selection behaviour on two versions of the same Google search engine results page (SERP), one showing advertisements and organic results, the other showing organic results only. Selection behaviour is analyzed in relation to users' knowledge on Google's business model, on SERP design, and on these users' actual performance in marking advertisements on SERPs correctly. We find that users who were not able to mark ads correctly selected ads significantly more often. This leads to the conclusion that ads need to be labeled more clearly, and that there is a need for more information literacy in search engine users.

Keywords: search engines; search engine results pages (SERPs); selection behaviour; representative user study; search engine advertising (SEA)

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 62–74.

1 Introduction

In the past years, we have been facing a discussion on the role search engines play for knowledge acquisition in society. The focus usually lies on Google, as it is by far the most dominant search engine in all countries except a few. Questions raised in this regard are, among others, what (types of) results users actually get to see when using Google (e.g., Noble, 2013; White & Horvitz, 2009), whether search engines are biased in general (Grimmelmann, 2010), and whether search engines in general and Google in particular should be made responsible for providing fair and unbiased results (Lewandowski, 2017).

Related to these questions, it is important to know how users actually interact with results presented by the search engines, first and foremost which results they select. The objective of the research presented in this paper is to find out whether users' knowledge on Google's business affects their clicking behaviour on the search engine results pages (SERPs). If so, i.e., if users who do not know that Google makes its revenue through advertisements on its results pages do select advertisements more frequently, then this would call for action either related to increasing users' information literacy and/or related to regulation.

To get reliable data and results regarding this, we conducted a large-scale online experiment, using a representative sample of the German online population.

The rest of this paper is structured as follows: In the next section, we give a brief overview on the elements presented on search engine results pages (SERPs), on selection behaviour on the SERPs, and on search engines' ad labeling practices. Then, we present our research questions and methods. After that, results are presented and discussed in a separate section each. We conclude with implications of our study and some suggestions for further research.

2 Literature review

In the following short literature review, we focus on the composition of search engine results pages, users' selection behaviour on these pages, and on ads labeling.

2.1 Search engine results pages

Results presentation on search engine results pages has changed in recent years. The simplest model of a search engine results page is a ranked list of document representations ("snippets") provided in response to a query. However, once search engines started displaying ads on SERPs, there were actually two ranked lists: the list of "organic results" and the list of advertisements. By adding results from vertical search engines such as news or video items and integrating them into the SERPs (known as Universal Search, cf. Taylor, Mayer & Buyukkokten, 2008), search engines moved away from plain ranked results lists and on to a richer presentation both of individual results as well as certain results types. Current search engine results pages go even further to additionally display factual information in addition to snippets. Factual information is shown in what are known as Knowledge Graph results (Drumond Monteiro & Aparecida Moura, 2014), satisfying at least some information needs directly on the results pages and representing a departure from the concept of a search engine being a tool for sending traffic (i.e., users) to external web pages.

For the purposes of the current study, we distinguish between the two results types *organic results* and *advertisements*. Other results types are omitted from the stimulus material.

2.2 Results selection

When looking at how users select results from a SERP, they are influenced mainly by the following factors:

1. Results position and reading behaviour: Users tend to click on results at or near the top of a results list (Joachims et al., 2005). For instance, a 2014 study based on 465,000 queries found that more than two thirds of all clicks go to the first five positions, and the result ranked first alone accounts for 31% of all clicks (Petrescu, 2014). Goel, Broder, Gabri-

lovich and Pang (2010) found that within Yahoo search, only 10,000 websites account for approximately 80% of results clicks. This clicking behaviour is due to the fact that, usually, lists of results are read from top to bottom.

2. Search engine relevance ranking algorithms are precision-based, i.e., they focus on presenting a few relevant results in the first several positions. Users have adapted to this kind of results ranking and therefore in most cases only consider the first few results (ibid.).
3. Due to screen resolutions and browser window sizes, SERPs can be divided into two areas, which are often referred to as “above the fold” vs. “below the fold” (Jansen & Spink, 2007), where the former refer to the results can be seen immediately without scrolling down. Users predominantly click on results shown above the fold. Therefore, it is important for content providers to take measures to make sure their content is listed in that area of the SERP.

2.3 Ad labeling

It is important to note that contextual, text-based ads (“sponsored links”, “paid results”) can be seen as one type of search result. They are similar to organic results in that they consist of a title, a short description and a URL, and they are also displayed on search engine results pages. Their design is also similar to organic results – the same colors or ones very similar to those used in organic results are used for headings, descriptions and URLs. Therefore, it seems reasonable to suppose that users may find it difficult to distinguish between the two results types.

Ads may be relevant to a query. What’s more, their uniqueness compared to other forms of advertisements lies precisely in the fact that a user has already entered a query and thereby expressed his or her intent (cf. Battelle, 2005).

An early study (Fallows, 2005) found that only 38% of U.S. searchers were aware of the distinction between organic and paid results. The situation has surely changed since then, partly due to the U.S. Federal Trade Commission’s guidelines on search engine ad disclosure (Sullivan, 2013a). However, the distinction between the two results types is still an issue, not only in general-purpose Web search engines, but also in many specialized vertical search engines (Sullivan, 2013b). Some industry studies strongly suggest that

the labeling of ads is not be clear enough (Bundesverband Digitale Wirtschaft, 2009; Charlton, 2013; Wall, 2012).

3 Research questions

The overall question guiding our research is whether users' knowledge on Google's business affects their clicking behaviour on the search engine results pages.

- *RQ1*: Does knowledge on the possibility of buying screen real estate on the SERPs (i.e., buying advertisements shown above the organic results) influence users' selection behaviour on the SERPs?
- *RQ2*: Does knowledge on the distinction between advertisements and organic results influence users' selection behaviour on the SERPs?
- *RQ3*: Does knowledge on Google's business model (i.e., selling ads) influence users' selection behaviour on the SERPs?
- *RQ4*: Does users' actual performance on distinguishing between organic results and paid advertisements on the SERPs influence their selection behaviour?

4 Methods

We conducted an online experiment using a representative sample of the German online population. The sample was built according to AGOF criteria ("Method – AGOF coverage currency", 2015) and consisted of 1,000 users. AGOF provides a standardized online coverage currency to measure the success of marketing tools. The online coverage currency is based on a Three-Pillar Model for Data mining and profiling by electronic measurement of page visits and page impressions, by on-site surveys on descriptive socio-demographic values and representative telephone surveys. The population includes Internet Users from the age of 10 years. In our study, each user was randomly assigned to either the experimental (ads) or control (no ads) condition.

Are Ads on Google Search Engine Results Pages Labeled Clearly Enough? 67

The stimulus material was two screenshots of a Google search engine results page (fig. 1). Both had the exact same results and the same layout, the only difference being one having the first two results labeled as ads (yellow shading, info button), while the other version only had a list of organic results. Note that we only used organic results, i.e., the ads shown in the experimental condition are actually organic results, only with an ads labeling.

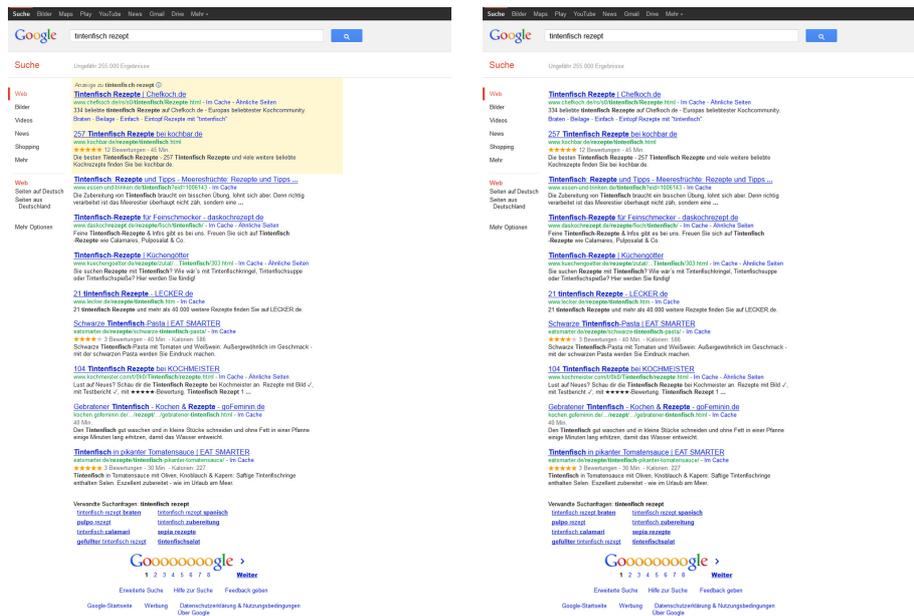


Fig. 1 Screenshots used as stimulus material

The participants were given the following search task: “Imagine you participate in a cooking competition where you should prepare fresh calamari. Your search query is ‘calamari recipe’. Which result(s) would you click on spontaneously?” [translated from German].

In our online experimental setting, participants were asked to mark the results they would select, i.e., we did not measure actual clicks but users labeled results as relevant. This allowed users to mark more than one result.

Each user was randomly assigned either to the experimental condition (ads) or the control condition (no ads). Additional data for the groupings used in the analysis below come from questions used in the data collection (detailed results on these can be found in Lewandowski et al., 2017). The questions relevant to the present study are the following:

- “Search engines are commercial Internet services, and therefore need to make money. Please describe in your own words how the search engine Google generates its revenues.” (This question operationalizes RQ3.)
- “Is it possible to pay Google for preferably listing one’s company on the search results pages, as an answer to a search query?” (This question operationalizes RQ1)
- “Is it possible to distinguish between paid advertisements and unpaid results on Google’s search engine results pages?” This question operationalized RQ2.

The fourth and last grouping (operationalizing RQ4) is built on four tasks where users had to mark all advertisements or all organic results on screenshots of SERPs. For detailed results on this task, see Lewandowski, Kerkmann, Rümmele and Sünkler (2017). For the present study, we only distinguish between users who proved in these tasks that they were able to identify ads consistently and those who were not able to do so.

5 Results

In the following, we analyze the clicks on the first two results (the ads in the experimental condition and the first two organic results in the control condition, respectively).

5.1 Knowledge-based questions

First, we looked at relationships between users’ knowledge on whether it is possible to buy screen real estate on the search engine results pages. There are three groups in this case: Those who say that it is possible (correct answer), those who say it is not possible (incorrect) and those how say they do not know. Looking at the selection behaviour in the two conditions, we do not find significant differences related to users’ knowledge (cf. table 1). While users in the control condition select the first two results more frequently than users in the experimental condition, there are no significant differences between groups within the conditions.

Table 1: Results to the question “Is it possible to pay Google for preferably listing one’s company on the search results pages, as an answer to a search query?”

Position	Experimental condition (ads)			Control condition (no ads)		
	Yes (n = 369)	No (n = 28)	Don’t know (n = 103)	Yes (n = 364)	No (n = 36)	Don’t know (n = 100)
1	38.5	42.9	36.9	59.1	63.9	55.0
2	25.7	17.9	33.0	38.7	41.7	44.0

* Differences between groups significant at $p \leq 0.05$

We found, however, significant results when it comes to users’ knowledge on whether it is possible to distinguish between paid advertisements and organic results on the search engine results pages (table 2). In the experimental group, users who say that this is possible select the first result significantly more often. However, the group saying they do not know select the top results even more often. As expected, there are no significant differences in the control condition.

Especially the result that users who say that it is possible to distinguish between the results types select the top result (in this case, an ad) more often seems counterintuitive. However, this may be explained by users who say this is not possible selecting the results they think are no advertisements.

Table 2: Results to the question “Is it possible to distinguish between paid advertisements and unpaid results on Google’s search engine results pages?”⁺

Position	Experimental condition (ads)			Control condition (no ads)		
	Yes (n = 217)	No (n = 90)	Don’t know (n = 62)	Yes (n = 208)	No (n = 105)	Don’t know (n = 51)
1	36.9*	32.2*	53.2*	59.6	54.3	66.7
2	23.5	23.3	37.1	37.0	38.1	47.1

* Differences between groups significant at $p \leq 0.05$

+ Only participant who answered “yes” in the preceding question (cf. table 1)

The third knowledge-based question related to users’ knowledge on how Google makes money from its search engine. We classified the answers into four groups: Correct answer (advertising), incorrect answer, partly correct answer (where advertising was mentioned, but other incorrect sources of revenue, as well), and “don’t know” (where users admitted they did not

know). We found significant differences in the selection behaviour of the different groups only in the control conditions (table 3). Users who know how Google makes money choose the first position significantly more often than users without that knowledge. A likely explanation is that these users notice that there are no ads on the page and therefore regard the first result as trustworthy.

Table 3: Results to the question “How does Google generate its revenue?”

Position	Experimental condition (ads)				Control condition (no ads)			
	Correct (<i>n</i> =300)	Incorrect (<i>n</i> =51)	Partly correct (<i>n</i> =100)	Don't know (<i>n</i> =49)	Correct (<i>n</i> =306)	Incorrect (<i>n</i> =42)	Partly correct (<i>n</i> =106)	Don't know (<i>n</i> =46)
1	40.0	35.3	38.0	32.7	61.1*	52.4*	39.1*	62.3*
2	27.3	21.6	25.0	32.7	43.5	42.9	40.4	33.0

* Differences between groups significant at $p \leq 0.05$

5.2 Performance measures

We found significant differences between selection behaviour on the first two results between users who proved to be able to distinguish between organic results and ads and those who were not able to do so (table 4). The latter chose advertisements significantly more often (40.3% vs. 21.6% for the first results, and 28.3% vs. 13.7%, respectively). The numbers are quite impressive: Users who are not able to distinguish between the two results types choose ads around twice as often as users who are able to recognize the ads. As expected, there are no significant differences in the selection behaviour in the control condition.

Table 4: Actual performance (marking ads)

Position	Experimental condition (ads)		Control condition (no ads)	
	All areas la- beled cor- rectly (<i>n</i> = 51)	Not all areas la- beled correctly (<i>n</i> = 449)	All areas la- beled correctly (<i>n</i> = 45)	Not all areas la- beled correctly (<i>n</i> = 455)
1	21.6*	40.3*	64.4	58.0
2	13.7*	28.3*	55.6	38.5

* Differences between groups significant at $p \leq 0.05$

6 Discussion

Our study revealed some surprising results, namely that users' self-reported knowledge on whether it is possible to pay for being shown on Google's SERPs did not affect their selection behaviour in the experimental condition (RQ1). We have to note, however, that we used *self-reported* measures, i.e., we do not know whether users saying that they know that it is possible to pay Google for being shown on the SERPs actually know how to spot paid results. Those who say it is possible to distinguish between ads and organic results select ads more often (RQ2). The likely reason is that they do so on purpose, as ads can be relevant to a search query (Jansen, 2007; Lewandowski et al., 2017). Users who know how Google generates its revenues selected the first results more frequently in the control condition (RQ3), which could be related to them trusting the search engine's ranking (Pan et al., 2007).

Maybe even more interesting, the performance measures based on distinguishing users into a group that actually *proved* to be able to distinguish between ads and organic results in a variety of tasks, and a group that was not able to do so, show a clear (and in some ways contradictory) result from the results discussed so far. They clicked on the ads approximately twice as often as the knowledgeable group. Through the experimental design of our study, we can rule out different relevance judgments as a reason for this, as users in both conditions saw exactly the same results.

Regarding methodology, the results of our study raise questions on the reliability of self-reported measures. Therefore, we argue for better using task-based user groupings than groupings based on questionnaires.

7 Conclusion

Our results have implications for search engine design and regulation, and for information literacy regarding search engines, as well. The results call for a clear labeling of advertisements on the search engine results pages. There has been a discussion on how ads should be labeled for at least 15 years now (Sullivan, 2013a), which has, however, not been based on a proper empirical basis. Our study contributes to filling this gap. However, it should be re-

garded as only a first step, as a limitation lies in that we used only one task, which was chosen because it is true-to-life and we assume all search engine users are able to choose relevant results to this task. However, further studies using tasks where the aim is to find high-quality information on professional tasks should be conducted.

Our results also call for more effort on helping users to become information literate when it comes to search engines. Efforts on increasing users' information literacy often focus on specialized information sources, and less on the tools that users are using on a daily basis. It is time for information literacy researchers and practitioners to focus on that blind spot.

Our study can only be a starting point on investigating users' selection behaviour when it comes to ads. Further research is needed, especially related to further results types, such as Universal Search results, which can be either ads or organic results.

References

- Battelle, J. (2005): *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*. London: Portfolio.
- Bundesverband Digitale Wirtschaft (2009): Nutzerverhalten auf Google-Suchergebnissen: Eine Eyetracking-Studie im Auftrag des Arbeitskreises Suchmaschinen-Marketing des Bundesverbandes Digitale Wirtschaft (BVDW). http://www.hottracker.de/whitepaper/Eyetrackingstudie_2008_Ergebnisbericht.pdf
- Charlton, G. (2013): 40% of consumers are unaware that Google Adwords are adverts. <https://econsultancy.com/blog/62249-40-of-consumers-are-unaware-that-google-adwords-are-adverts>
- Drumond Monteiro, S., and M. Aparecida Moura (2014): Knowledge Graph and "Semantization" in Cyberspace: A Study of Contemporary Indexes. In: *Knowledge Organization*, 41 (6), 429–439.
- Fallows, D. (2005): Search engine users: Internet searchers are confident, satisfied and trusting – but they are also unaware and naive. Pew Internet & American Life Project. Washington, DC: Pew Internet & American Life Project.
- Goel, S., A. Broder, E. Gabrilovich, and B. Pang (2010): Anatomy of the long tail: Ordinary people with extraordinary tastes. In: *Proceedings of the third ACM international conference on Web search and data mining* (pp. 201–210). ACM.

Are Ads on Google Search Engine Results Pages Labeled Clearly Enough? 73

- Grimmelmann, J. (2010): Some Skepticism About Search Neutrality. In: *The Next Digital Decade: Essays on the Future of the Internet*, 435–459.
- Jansen, B. J. (2007): The comparative effectiveness of sponsored and nonsponsored links for Web e-commerce queries. In: *ACM Transactions on the Web*, 1 (1), article 3.
- Jansen, B. J., and A. Spink (2007): Sponsored search: is money a motivator for providing relevant results? In: *Computer*, 40 (8), 52–57.
- Joachims, T., L. Granka, B. Pan, H. Hembrooke, and G. Gay (2005): Accurately interpreting clickthrough data as implicit feedback. In: *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 154–161). Salvador, Brazil: ACM.
- Lewandowski, D. (2017): Is Google responsible for providing fair and unbiased results? In: L. Floridi and M. Taddeo (Eds.): *The Responsibilities of Online Service Providers*. Berlin, Heidelberg: Springer.
- Lewandowski, D., F. Kerkmann, S. Rümmele, and S. Sünkler (2017): Search engine ad disclosures. [forthcoming]
- Method – Method – AGOF coverage currency (2015). <https://www.agof.de/studien/internet-facts/methode-internet/?lang=en>
- Noble, S. U. (2013): Google Search: Hyper-visibility as a Means of Rendering Black Women and Girls Invisible. <http://ivc.lib.rochester.edu/google-search-hyper-visibility-as-a-means-of-rendering-black-women-and-girls-invisible/>
- Pan, B., H. Hembrooke, T. Joachims, L. Lorigo, G. Gay, and L. Granka (2007): In Google we trust: users' decisions on rank, position, and relevance. In: *Journal of Computer-Mediated Communication*, 12 (3), 801–823.
- Petrescu, P. (2014): Google Organic Click-Through Rates in 2014. <https://moz.com/blog/google-organic-click-through-rates-in-2014>
- Sullivan, D. (2013a): FTC Updates Search Engine Ad Disclosure Guidelines After “Decline in Compliance.” <http://searchengineland.com/ftc-search-engine-disclosure-164722>
- Sullivan, D. (2013b): Revealed: The 17 Other Search Engines The FTC Warned Over Paid Ad Disclosures. <http://searchengineland.com/revealed-17-other-search-engines-ftc-warned-168603>
- Taylor, B., M. Mayer, and O. Buyukkokten (2008): Interface for a universal search engine. United States Patent and Trademark Office.
- Wall, A. (2012): Consumer Ad Awareness in Search Results. <http://www.seo-book.com/consumer-ad-awareness-search-results>

White, R. W., and E. Horvitz (2009): Cyberchondria. In: *ACM Transactions on Information Systems*, 27 (4): Article No. 23. <http://doi.org/10.1145/1629096.1629101>

Session 2:

Designing Scholarly Information Systems

A Reference Architecture for Virtual Research Environments

*Keith G. Jeffery¹, Carlo Meghini², Cesare Concordia²,
Theodore Patkos³, Valerie Brasse⁴, Jacco van Ossenbruck⁵,
Yannis Marketakis³, Nikos Minadakis³, Eda Marchetti²*

¹ Keith G Jeffery Consultants, UK
keith.jeffery@keithgjefferyconsultants.co.uk

² ISTI – CNR, Italy
{carlo.meghini, cesare.concordia, eda.marchetti}@isti.cnr.it

³ Information Systems Laboratory, FORTH, Greece
{patkos, marketak, minadakn}@ics.forth.gr

⁴ euroCRIS / IS4RI, France
vbrasse@is4ri.com

⁵ CWI, Amsterdam, The Netherlands
Jacco.van.Ossenbrugge@cwi.nl

Abstract

This paper describes the Reference Architecture of the *enhanced VRE* (e-VRE), a Virtual Research Environment defined in the context of the VRE4EIC Project, funded by EU H2020 e-Infrastructure program. e-VRE is designed to overcome limits of existing VREs with respect to a number of orthogonal dimensions: improving the quality of VRE user experience by providing user centered, secure, privacy compliant, sustainable environments for accessing data, composing workflows and tracking data publications; increasing VRE usage in multidisciplinary research domains by abstracting and reusing building blocks and workflows from existing VRE initiatives; improving the interoperability of heterogeneous discovery, contextual and detailed metadata across all layers of the VRE.

Keywords: Virtual Research Environments; multidisciplinary; interoperability; innovation; collaboration; distributed systems architecture; use cases

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 76–88.

1 Introduction

The goal of a VRE system is to decouple Science from ICT complexity, by providing researchers with a facility that takes care of ICT so allowing them to focus on their work. In this sense, a VRE is a fundamental component running on top of an e-Research Infrastructure (e-RI) (Carusi & Torseten, 2010) as it purports at making the resources of the e-RI easily accessible and reusable to the community of researchers that owns the e-RI.¹ Here, by e-RI we mean “facilities, resources and related services used by the scientific community to conduct top-level research in their respective fields”², while *resource* indicates any ICT entity that is of interest in an e-science community. Typically, a resource is owned by an e-RI that provides an identity for the resource and manages it, making it accessible and reusable. Examples of resources are: datasets, workflows, algorithms, Web Services, computational or storage facilities, cloud endpoints etc.

In general, a VRE is expected to:

- allow researchers to communicate with each other and to use and share the resources available in the community’s e-RI;
- allow researchers to advance the state of the art by building new resources as the result of processing existing resources with the available tools. Such processing may be the application of an individual piece of software to a dataset, such as the extraction of certain knowledge from a single file; or, it may result from the execution of a complex workflow combining available services, including other workflows, on a number of data resources;
- allow research managers to apply economy of scale models to access and manage resources that researchers or single organizations alone could not afford.

Moreover, a VRE can offer all of the above on top of an individual e-RI or on top of several e-RIs, the latter option clearly requiring a level of interoperability that would empower researchers and managers in ways that are only imaginable today.

The most advanced e-RIs have developed their own VRE, showing awareness of the crucial role that a VRE can play for their researchers. Others are currently designing their VRE. However, the number of currently existing or

1 <https://www.jisc.ac.uk/full-guide/implementing-a-virtual-research-environment-vre>

2 http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=what

designed VREs is very limited; more importantly, these VREs show a great heterogeneity in scope, features, underlying protocols and technologies, partially defeating the interoperability goal that lies at the very heart of a VRE.

VRE4EIC (*A Europe-wide Interoperable Virtual Research Environment to Empower Multidisciplinary Research Communities and Accelerate Innovation and Collaboration*) is a three-year project funded by the European Commission in the context of the H2020 Program, under the topic EINFRA-9-2015 *e-Infrastructures for virtual research environments*. One of the major goals of the VRE4EIC project is to overcome the above described issues by building an *enhanced VRE* (e-VRE) system whose main features are:

- Increase the quality of VRE User Experiences (UX) by providing user centered, secure, privacy compliant, sustainable environments on searching data, composing workflows and tracking data publications.
- Increase the deployment of the VRE on different clusters of research infrastructures by abstracting and reusing building blocks and workflows from existing VREs, infrastructures and projects.
- Improve the contextual awareness and interoperability of the metadata across all layers of the resources in the VRE.
- Promote the exploitation and standardisation of the VRE4EIC solution to different research domains and communities.
- Provide interoperation across ‘silo’ e-RIs.

The main step to implement the above features is to create a Reference Architecture that can serve as a guide for the development of enhanced VREs. This paper describes the Reference Architecture produced by the VRE4EIC project after its first year of activity. The paper is structured as follows: next section presents related work in the same area, while the following section outlines the methodology that has inspired our work and the way this methodology has been adapted to the present context. Next, the paper presents the vision of a VRE that lies at the basis of our Reference Architecture, presented in the following section. The paper then concludes after briefly alluding at the further development of the Architecture within the VRE4EIC project.

2 Related work

In designing the reference Architecture of e-VRE we have taken cognizance of past and ongoing work on:

- Science Gateways (SG). In general SGs are portals to datasets. Some have analytical, visualisation and simulation capabilities. Some provide access to – and steering of – equipment. Some provide access to specialist computing resources and some provide collaboration tools. However – in contrast to VRE4EIC – in general they are constructed on top of one or more – e-RIs in a particular domain.
- Virtual Laboratories (VL). VLs have taken a different approach, they are constructed from a pool of general software modules, available datasets and user groups. Again each VL tends to be domain specific and linked with one or a small number of e-RIs.

The booklet produced by DG-CNECT on Research Infrastructures is a useful reference resource.³ Additionally the VRE4EIC team is carefully checking and tracking activities related to other EU projects on VRE. Indeed, the VRE4EIC project is one of the four H2020 RIA projects concerned with Virtual Research Environments; other projects are: EVER-EST⁴ (geoscience); BlueBridge⁵ (dominantly marine) and West-LIFE⁶ (Bio). All projects are currently at early stage of development, so our considerations have relied on (a) information from the project websites (b) personal contacts especially with EVER-EST and Blue Bridge (where the major partner is the same organization, but a different group, as the architecture developers in VRE4EIC). The significant differences in approaches are:

1. VRE4EIC is producing a reference architecture (and prototype demonstrator) that can bridge across e-RIs (and hence underlying e-Is) in a multidisciplinary manner; the other projects are restricted to particular domains;
2. BlueBridge produces a VRE that is tightly coupled to the underlying e-RIs;

3 <https://ec.europa.eu/digital-single-market/en/news/e-infrastructures-making-europe-best-place-research-and-innovation>

4 <http://www.ever-est.eu>

5 <http://www.bluebridge-vres.eu>

6 <https://portal.west-life.eu>

3. EVER-EST is using research objects; this binds data and code in a particular way that restricts openness and interoperability, which is unacceptable for the objectives of VRE4EIC. Nonetheless we are working together and believe there are opportunities for co-development.
4. WEST-LIFE is built in context of the European Grid Initiative and its modularized components for self-assembly by the e-RIs to form a VRE.

Outside of these H2020 projects for VRE4EIC participants, the scientific coordinator has initiated a RDA IG (Research Data Alliance Interest Group) jointly with EVER-EST.

3 Methodology and approach

In the development of the Reference Architecture, an incremental software development process has been adopted, largely inspired by the RUP process⁷ (Jacobson et al., 1999). Architectural components are derived based on an analysis of the requirements collected in the project. Figure 1 presents a UML class diagram outlining the entities involved in this analysis and their relationships. In particular:

- The analysis started from the Requirements (yellow box in fig. 1). Each requirement has been considered individually, and the functions (green box) required for its implementation have been derived.
- In order to ease the specification of functions, a set of generalised functions has also been derived, which are included or specialised by functions, or which may be used as preconditions by functions.
- The components that are required for the implementation of functions have finally been derived.

As the figure also shows, this analysis of requirements into functions and components connects to the use cases (turquoise box) and the high-level use cases (purple box). The connection is realized through requirements, and is used for the assessment of the Reference Architecture. Overall, the schema shown in figure 1 allows us to maintain the relationship between use cases, requirements and components, thereby realizing the traceability of the Reference Architecture.

⁷ https://www.ibm.com/developerworks/rational/library/content/03July/1000/1251/1251_bestpractices_TP026B.pdf

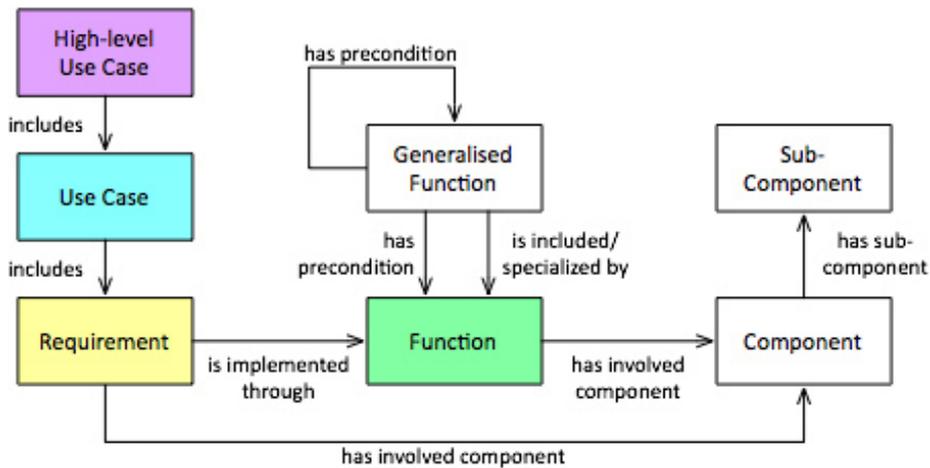


Fig. 1 Entities and relationships involved in the analysis of requirements

4 A vision for the e-VRE architecture

In order to enable a VRE to make its resources available to the researchers that use the VRE, each e-RI that *participates* to a VRE must provide *descriptions* of its resources, and such descriptions must be rich enough in information to support the VRE services. This information may include the protocol that must be used to interact with an e-RI service; the schema, size and operations allowed on an e-RI dataset; the permission framework that must be adopted for authentication/authorization of the e-RI users, and so on. This process is naturally divided into two steps, as depicted by figure 2: the e-RI resources are given at the bottom level since they are the basic assets that both e-RIs and VREs operate on; at the next level up, the descriptions of these resources used by the e-RI services (*e-RI Resource descriptions*) are given, next to the e-RI services using them; at the top level, the descriptions of the e-RI resources used by the VRE services (*VRE Resource descriptions*) are given, next to the VRE services using them. In both cases, the services mentioned are purely exemplificative.

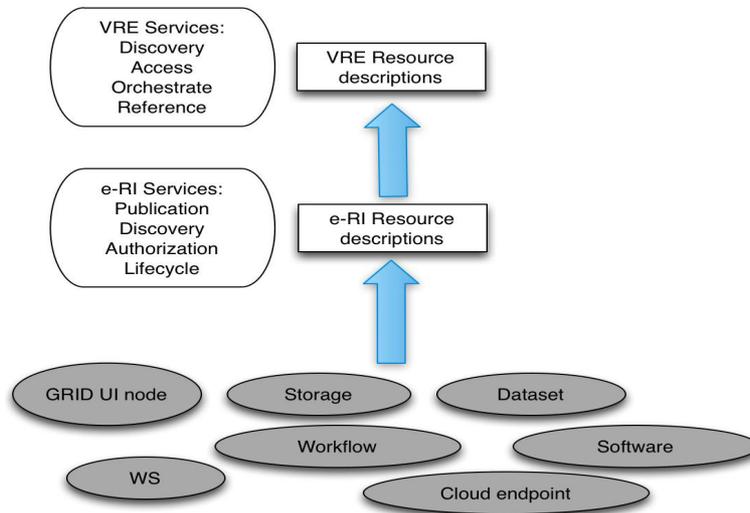


Fig. 2 Resources, e-RIs Services and VRE Services

Resource descriptions are typically collected in *Catalogues*. So, the VRE needs to access the catalogues of the participating e-RIs in order to discover the existing resources in the e-RI and obtain enough information on these resources to create its own descriptions of them in its own catalogue. In order to simplify the creation of the VRE catalogue, one could employ the same data model for both the e-RI and the VRE descriptions. In fact, VREs that live within e-RIs follow this approach. In this case, e-RIs and VRE Resource descriptions only differ for the type of information they contain, while sharing the identity and the basic attributes of resource descriptions. However, this approach is not feasible for VREs with *many* participating e-RIs, due to the fact that in general different e-RIs use different data models to structure their catalogues. In this case, there are two main approaches to create and maintain the VRE catalogue:

- The *centralized* approach (cf. fig. 3, left), in which there exists a VRE Catalogue used by the VRE services for carrying out their own operations.
- The *distributed* approach (cf. fig. 3, right), in which there is no VRE Catalogue, but the VRE accesses the e-RIs catalogues when the information is needed.

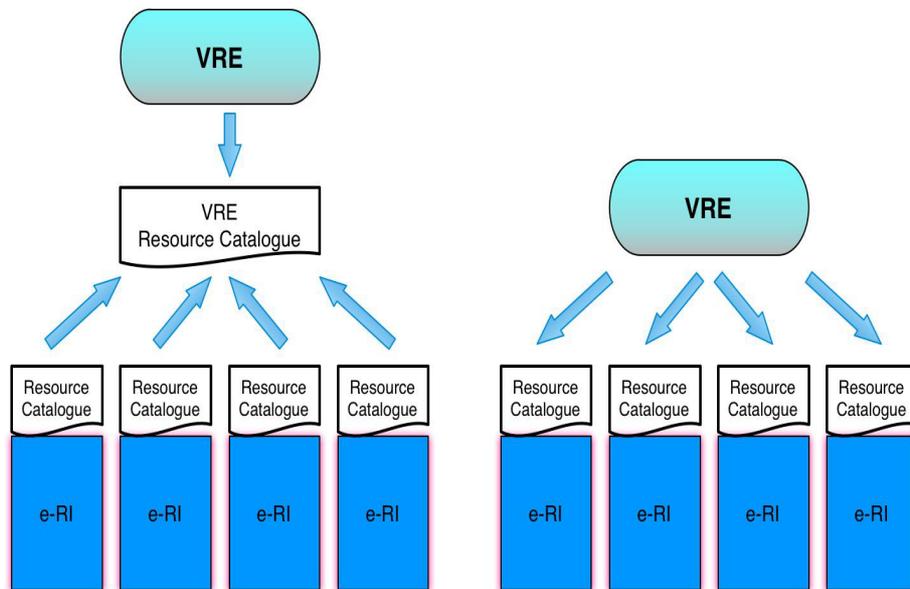


Fig. 3 Alternative approaches to the cooperation on e-RIs and a VRE

Each approach has its own pros and cons, as it is well known in distributed system design. In fact, the availability of a VRE Catalogue facilitates all VRE operations that rely exclusively on resource descriptions, such as resource discovery. For operations that require data access, such as data discovery, the centralized approach can only alleviate the problem, by offering information for executing part of the operation. On the other hand, the distributed approach makes it easier to have complete information in real time, since it does not require propagation of updates to the Catalogue. Our Reference Architecture chooses the centralized approach, because it facilitates one important service, namely the construction of workflows across one or more RIs. The construction of workflows requires numerous access to resource descriptions, followed by optimisation for parallel/distributed operations; the centralized approach makes it possible to implement this access in the most efficient way possible.

5 A reference architecture for e-VRE

At a more general level, the Reference Architecture conforms to the multi-tiers view paradigm used in the design of distributed information systems. Following this paradigm, we can individuate three logical tiers in e-VRE, as shown in figure 4:

- The *Application* tier, which provides functionalities to manage the system, to operate on it, and to *expand* it, by enabling administrators to plug new tools and services into the e-VRE.
- The *Interoperability* tier, which deals with interoperability aspects by providing functionalities for: i) enabling application components to discover, access and use e-VRE resources independently from their location, data model and interaction protocol; ii) publishing e-VRE functionalities via a Web Service API; and iii) enabling e-VRE applications to interact each others.
- The *Resource Access* tier, which implements functionalities that enable e-VRE components to interact with eRIs resources. It provides synchronous and asynchronous communication facilities.

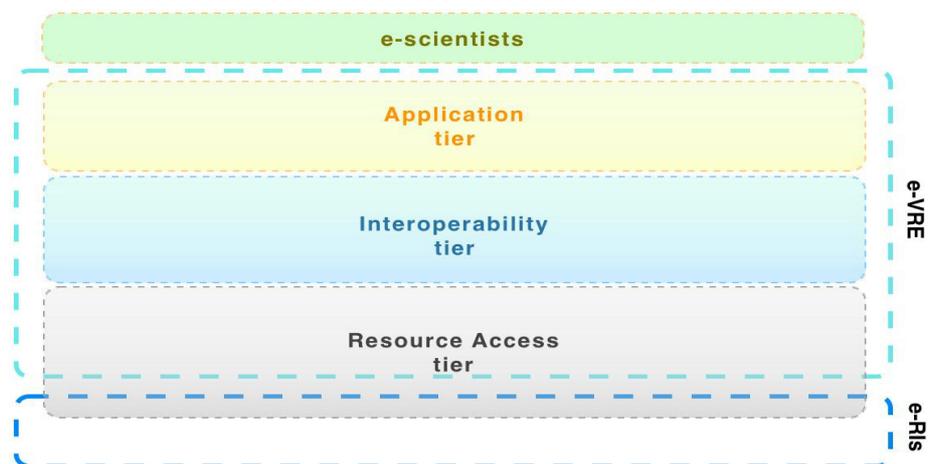


Fig. 4 Architectural tiers in a VRE

Figure 4 depicts the logical tiers of e-VRE and shows their placement in an ideal space between the e-scientists that use the e-VRE and the e-RIs that provide the basic resources to the e-VRE. Generally speaking a VRE system can be viewed as a dynamic framework; it “is the result of joining together

new and existing components to support as much of the research process as appropriate for any given activity or role”⁸. To implement this fundamental non-functional requirement the e-VRE system has been designed following a *component-oriented* approach. According to this approach a system is composed by an integration infrastructure where a set of software components can be deployed, these components implement the system functionalities and potentially can be specified, developed and deployed independently of one another. Based on these considerations and on the analysis of the requirements, for the basic integration infrastructure of e-VRE we have individuated a set of basic functionalities grouped into six *conceptual components*:

- The e-VRE management is implemented in the **System Manager** component. The System Manager can be viewed as the component enabling Users to use the *core* functionalities of the e-VRE: access, create and manage resource descriptions, query the e-VRE information space, configure the e-VRE, plug and deploy new tools in the e-VRE and more.
- The **Workflow Manager** enables users to create, execute and store business processes and scientific workflows.
- The **Linked Data (LD) Manager** is the component that uses the LOD (Linked Open Data) paradigm, based on the RDF (Resource Description Framework) data model, to publish the e-VRE information space – i.e. the metadata concerning the e-VRE and the e-RIs in a form suitable for end-user browsing in a SM (Semantic Web)-enabled ecosystem.
- The **Metadata Manager** (MM) is the component responsible for storing and managing resource catalogues, user profiles, provenance information, preservation metadata used by all the components. All these entities and their relations are captured in the CERIF data model⁹ (Jeffery et al., 2002), using extended entity-relational conceptual and object-relational logical representation for efficiency.
- The **Interoperability Manager** provides functionalities to implement interactions with e-RIs resources in a transparent way. It can be viewed as the interface of e-VRE towards e-RIs. It implements services and algorithms to enable e-VRE to: communicate synchronously or asynchronously with e-RIs resources, query the e-RIs catalogues and storages, and map the data models.

⁸ Fraser, M.: “Virtual Research Environments: Overview and Activity”, 30-July-2005, <http://www.ariadne.ac.uk/issue44/fraser/>

⁹ <http://eurocris.org/ontologies/semcerif/1.3/>

- The **Authentication, Authorization, Accounting Infrastructure (AAAI)** component is the responsible for managing the security issues of the e-VRE system. It provides user authentication for the VRE and connected e-RIs, authorisation and accounting services, and data encryption layers for components that are accessible over potentially insecure networks.

Figure 5 shows how these six components are distributed on the 3-tier space introduced above.

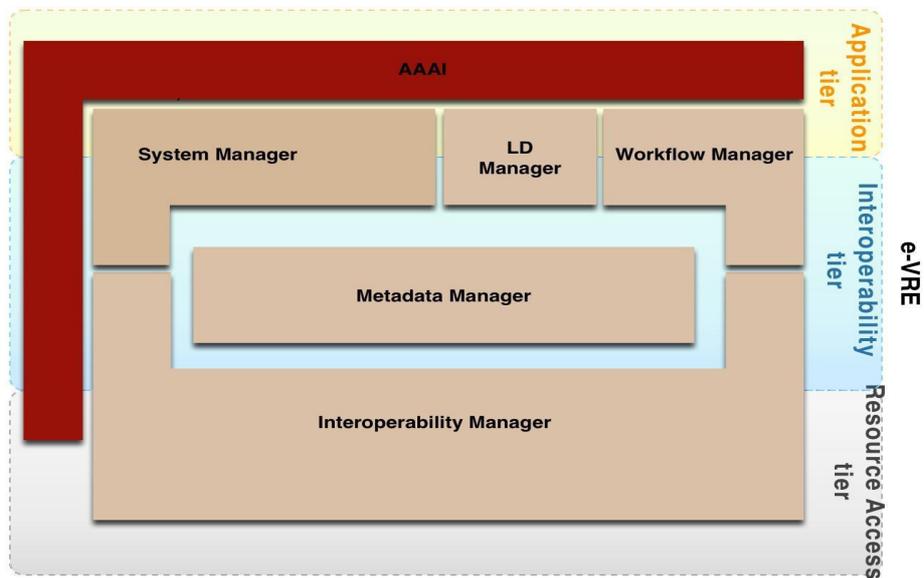


Fig. 5 Conceptual components and logical tiers

6 Outlook

A first version of the Reference Architecture has been released at the end of the first year of the project, that is end of September 2016. The Reference Architecture has been provided in terms of two kinds of UML diagrams: a component diagram highlighting the components of the architecture, their provided and used interfaces; and a number of sequence diagrams highlighting the interactions occurring in the execution of the main methods. A revised version of the first release will be produced by the end of the second year of the project (end of September 2017), for internal usage only. This

revision has been deemed as necessary by the project management as a consequence of the fact that the list of user requirements will be in turn refined three times during the second year of the project. These refinements have been planned in order to accommodate the collection of a large amount of requirements, along with the ongoing characterization of existing e-RIs. Each time the requirements and the e-RI characterizations will be updated, the Reference Architecture will consequently be revised. In parallel to the revision of the architecture, a Gap Analysis work will be performed, to determine the components that will be implemented. The development process will rely on the re-use of existing technologies and standards, which in turn may lead to the revision of some interfaces of the Reference Architecture, for instance to align an interface with the selected standard or technology; this alignment may be propagated into the Reference Architecture, if the standard or technology provoking it are important enough.

7 Conclusions

The Reference Architecture for Virtual Research Environments developed in the context of the VRE4EIC Project has been presented, based on a vision of VREs and of the relationships between VREs and their close relatives, i.e. e-Research Infrastructures, e-Infrastructures, Science Gateways and Virtual Laboratories. The Architecture has been described in terms of its main components, organized in a three-layered approach that closely resembles classic three-tier architectures. These components have been derived by analysing a large set of requirements, collected in more than 40 interviews, and by characterising 5 existing e-Research Infrastructures [D2.1]. A complete specification of the Reference Architecture, as well as a detailed documentation of the process that led to its derivation, can be found in the deliverable D3.1 of the VRE4EIC Project [D3.1], which presents the main UML component diagram of the architecture, highlighting the interfaces provided and used by each component. Such interfaces are further specified in terms of the signatures of the methods that are included in each of them. UML sequence diagrams are also provided for the main methods and use cases. The flow of work within the VRE4EIC Project that will lead to the refinement, validation, and partial implementation of the Reference Architecture has also been presented, put-

ting in context the present Architecture and indicating its evolution in the next two years of work.

Acknowledgements

The research leading to these results has received funding from the European Union Horizon 2020 Programme, Topic: e-Infrastructures for virtual research environments, Research and Innovation action VRE4EIC (A Europe-wide Interoperable Virtual Research Environment to Empower Multidisciplinary Research Communities and Accelerate Innovation and Collaboration), grant agreement n 676247.

References

- Carusi, Annamaria and Torsten Reimer (2010): Virtual Research Environment. Collaborative Landscape Study. <http://www.webarchive.org.uk/wayback/archive/20140614205957/http://www.jisc.ac.uk/publications/reports/2010/vrelandscapestudy.aspx>
- Jacobson, I., G. Booch, J. Rumbaugh, and Booch, G. (1999): *The unified software development process* (Vol. 1). Reading: Addison-Wesley.
- Jeffery, Keith G., Andrei Lopatenko and Anne Asserson (2002): Comparative Study of Metadata for Scientific Information: The place of CERIF in CRISs and Scientific Repositories. In: *6th International Conference on Current Research Information Systems, Kassel, Germany*.

VirtualPatent

Enabling the Traceability of Ideas Shared Online Using Decentralized Trusted Timestamping

Corinna Breitinger, Bela Gipp

Department of Computer and Information Science
University of Konstanz, Germany
{corinna.breitinger, bela.gipp}@uni-konstanz.de

Abstract

Online discussion and sharing platforms have enabled ideas to be disseminated more quickly than ever before. However, there are many good reasons why individuals hesitate to share their ideas online. In academia, for instance, researchers may not want their contribution to be made public until after it has been published to ensure that they are appropriately credited for their work. As a consequence, novel ideas or creative work tend to only be shared within a small circle of trusted peers instead of with wider audiences online. This status quo prevents other experts on a specific topic from contributing to the discussion. In this paper, we present a proof-of-concept implementation of an online discussion and sharing platform that addresses this problem. The web-based application, coined VirtualPatent, automatically timestamps each post a user shares by creating a distributed timestamp on the blockchain of the cryptocurrency Bitcoin – a method for trusted-timestamp creation that we published in a previous paper. Unlike platform-managed timestamps, timestamps stored on the blockchain are persistent and cannot be tampered with. The system thus enables the author of a posting made online to retrospectively prove the exact time that the specific contribution was first put forth in a tamperproof manner – similar to a published paper, but with the simplicity of writing a post on a social media website.

Keywords: sharing platforms; social media, trusted timestamping; intellectual property; blockchain applications

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 89–95.

1 Introduction¹

The volume of content shared online is growing at unprecedented rates. As of 2016, every 60 seconds: 347,000 tweets are published on Twitter, 66,000 photos and videos uploaded to Instagram,² and 400 hours of video content added to YouTube (James, 2015). Unlike physical media, digital contributions online are susceptible to widespread distribution and modification, with no reliable mechanism for tracking content back to its earliest point and time of origin. This volatility of Web resources demands standards to “assure us that information can be verified and traced to its source” (Snapper, 1999: 127). Today, commercial services, such as CopyScape³ or Searchlight⁴, offer to search the Web for unauthorized redistribution of ideas and creative content. However, no tamperproof mechanism has been integrated directly into a social media platform to enable content creators, such as authors, academics, artists, musicians, photographers, or innovators to prove priority for their work and ideas when shared online.

Existing social discussion and sharing platforms typically associate all user-uploaded media with a timestamp, however, such platform-generated timestamps share the same shortcomings: They are (a) not tamperproof, and (b) not persistent. Centrally managed media timestamps can be manipulated, e.g. by the platform operator or hacked by an external party. Even when no malice is involved, social sharing platforms can experience technical failures, or simply cease to exist, resulting in no guarantee for platform-managed media timestamps to be accessible indefinitely.

In this paper, we propose attaching a mark of traceability and permanence to user contributions online using trusted timestamping on the blockchain. We subsequently demonstrate the approach in a proof-of-concept web application.

1 This paper includes research performed in the scope of the first author’s Master thesis (Breitinger, 2016).

2 <https://www.instagram.com/press/>

3 <http://www.copyscape.com>

4 <http://www.searchlig.ht>

2 Trusted timestamping of digital media

Decentralized Trusted Timestamping (DTT) using the blockchain is a recently introduced technical solution to solve the problem of securely verifying the time at which digital content existed in a certain state (Gipp, Meuschke & Gernandt, 2015). In this digital timestamping approach, the unique hash digest, e.g. SHA-256, generated from the given digital file is embedded as a transaction in a cryptocurrency's decentralized blockchain.⁵ Due to the infrastructure of Bitcoin's blockchain (Nakamoto, 2008), it is computationally infeasible to manipulate transaction records, meaning timestamps are stored on a tamperproof and persistently verifiable medium.

While timestamping digital files is not new (Haber & Stornetta, 1991), DTT is an improvement upon the traditional digital timestamping (DTS) protocols that rely on a centralized third party to act as a Time Stamping Authority (TSA). Timestamping protocols include the RFC 3161 standard for trusted timestamps (Adams & Pinkas, 2001), or the ANSI ASC x9.95 Standard (ANSI, 2005). Unlike DTS, DTT does not require users to place trust in a TSA, which are typically commercial, and thus eliminates the risk of generated timestamps becoming invalid, e.g. if the TSA's private keys used for the public key encryption scheme were to become compromised.

Today, few content creators benefit from trusted timestamping to secure the time of existence of their contributions. We present an easy-to-use platform that implements DTT in the background to timestamp user-uploaded digital media, free of charge and with no added effort. We hypothesize that providing a method for securely verifying the date of existence of media shared online – independent of the distribution platform – is a key step toward (1) making the origin of ideas traceable and (2) increasing the user's incentive to share.

⁵ Space limitations prohibit us from discussing blockchain in more detail; for an overview of blockchain technology and its implications, please consult (Swan, 2015).

3 System design and implementation

The system is implemented as a *flask*⁶ web application with a Python backend and a SQLite database. The application is hosted on Heroku as a demo and should be seen as a proof-of-concept⁷. We invite others to develop their own applications upon our idea, which is why the source code is available under an MIT license at: www.gipp.com/dtt. On the platform, users can create profiles and upload a variety of file formats. In a previous paper, we introduced a mobile application that focused exclusively on timestamping video files, specifically the video recorded by dashboard cameras (Gipp, Kosti & Breitingner, 2016). VirtualPatent calculates the unique hash for the uploaded file using *hashlib*'s SHA-256 function⁸. This hash is submitted for timestamping in Bitcoin's blockchain via OriginStamp's API.⁹ Figure 1 shows the communication between the web application and the OriginStamp service to generate the timestamp on the blockchain. To avoid bloating the blockchain with unnecessary transactions, and to minimize transaction costs (to keep the timestamping service free of charge), OriginStamp collects all hashes received over a 24-hour period and generates one aggregate hash (SHA-256), see top right box in figure 1. From this aggregate hash, a new Bitcoin address is computed to which a one-time transaction of a Satoshi (0.00000001 Bitcoin) is sent. The transaction time at which the hash is embedded into the blockchain network becomes the timestamp of the file. Due to space limitations, please consult (Gipp et al., 2015) for the technical details on timestamp generation.

6 <http://flask.pocoo.org/>

7 The demo is accessible via: <https://www.gipp.com/virtualpatent>.

8 <https://docs.python.org/3/library/hashlib.html>

9 <http://www.originstamp.org/developer>

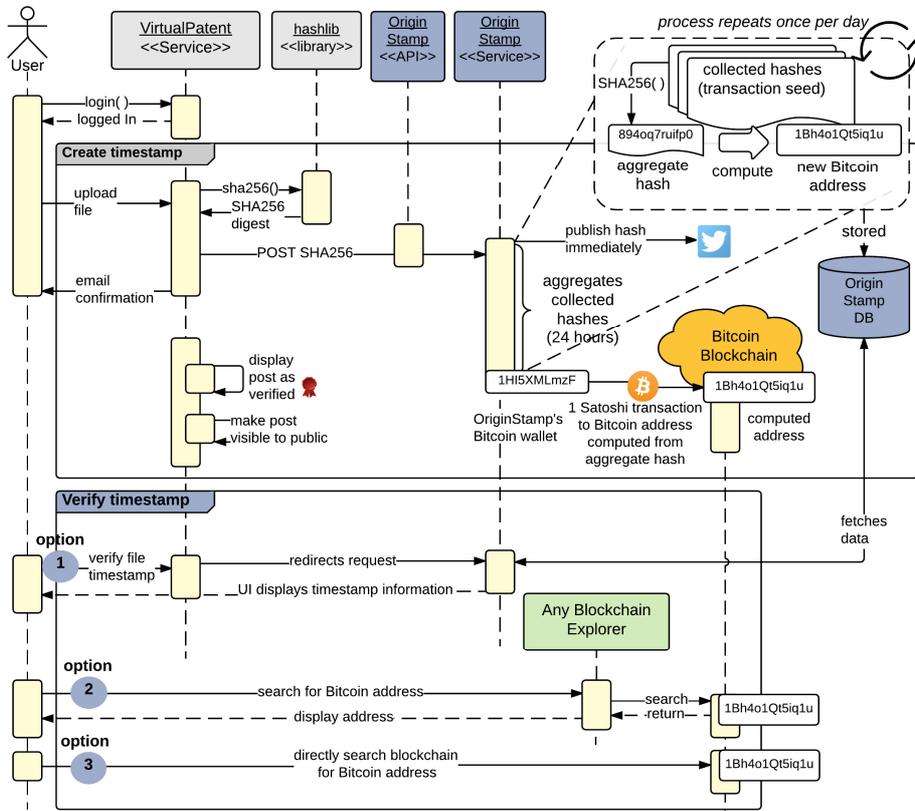


Fig. 1 Timestamp creation and verification process for files shared on VirtualPatent

Since the content is timestamped with a delay, users have the choice to keep their post private until the transaction has been confirmed in the blockchain network. To verify the timestamp of a file, the unique file hash and a ‘verification link’ are displayed with each post. The verification link currently redirects to OriginStamp’s database of submitted timestamps, but could redirect to any blockchain explorer that provides users a GUI for verifying transactions in the blockchain network, see ‘option 2’ in figure 1. VirtualPatent additionally sends registered users an email confirmation once their file has been timestamped. The email includes the list of hashes used to generate the aggregate hash from which the Bitcoin address was computed. Sharing this data with the user guarantees that even if our platform or the OriginStamp service ceases to exist, users have a copy of all data required to verify that their file’s hash was indeed embedded in a transaction on the blockchain.

The frontend of the VirtualPatent platform emphasizes ease of use and supports the actions that users expect of online discussion and sharing platforms, including profile creation and management, browsing posts, liking and sharing posts, commenting on posts, and following users. The system's functionality was inspired by the widely-used idea and knowledge exchanging platforms Quora and StackExchange.

4 Conclusion

Discussion and sharing platforms have enabled the barrier-free dissemination of creative work and ideas. However, no reliable method exists for tracing original ideas to their first time of existence online. We proposed the automatic and immutable timestamping of digital media shared online using trusted timestamping on the blockchain and presented VirtualPatent, a proof-of-concept platform. The timestamps that VirtualPatent creates for users' files are independent of the social sharing and discussion platform itself. Even if the platform were to cease to exist, the timestamps associated with the files can still be verified in the decentralized ledger that is the blockchain. This allows the user to prove the date of origin for their idea or contribution. Having an accessible and secure means to prove priority for an idea may encourage creatives to more openly share ideas and work online. In the future, we plan to integrate the approach demonstrated by VirtualPatent into other document management tools, such as Docear (Beel et al., 2011; Beel et al., 2014), to enable researchers to timestamp their ideas while writing their manuscript.

Acknowledgements

This research was partially funded by the Carl-Zeiss Foundation.

References

- Adams, C., and D. Pinkas (2001): Internet X. 509 public key infrastructure time-stamp protocol (TSP).
- American National Standards Institute (ANSI) (2005): X9 x9.95-2005 trusted time stamp management and security.
- Beel, J., B. Gipp, S. Langer, and M. Genzmehr (2011): Docear: An Academic Literature Suite for Searching, Organizing and Creating Academic Literature. In: *Proceedings of the 11th ACM/IEEE Joint Conference on Digital Libraries (JCDL '11)*. ACM. <http://doi.org/10.1145/1998076.1998188>
- Beel, J., S. Langer, B. Gipp, and A. Nuernberger (2014): The Architecture and Datasets of Docear's Research Paper Recommender System. In: *D-Lib Magazine – The Magazine of Digital Library Research*, 20 (11/12). <http://www.dlib.org/dlib/november14/beel/11beel.html>
- Breitinger, C. (2016): *Using the Blockchain of Cryptocurrencies to Encourage Open Discussion and Sharing of Ideas*. MS Thesis, Linnaeus University.
- Gipp, B., J. Kosti, and C. Breitinger (2016): Securing Video Integrity Using Decentralized Trusted Timestamping on the Blockchain. In: *Proceedings of the 10th Mediterranean Conference on Information Systems (MCIS)*, Paphos, Cyprus.
- Gipp, B., N. Meuschke, and A. Gernandt (2015): Decentralized Trusted Timestamping using the Crypto Currency Bitcoin. In: *Proceedings of the iConference 2015*. Newport Beach, California. <http://ischools.org/the-icconference/>
- Haber, S., W.S. Stornetta (1991): How to Time-Stamp a Digital Document. In: *Advances in Cryptology – CRYPTO '90 Proceedings*, 3 (2), 99–111. <http://doi.org/10.1007/BF00196791>
- James, J. (2015): Data Never Sleeps 3.0. <https://www.domo.com/blog/data-never-sleeps-3-0/>
- Nakamoto, S. (2008): Bitcoin: A peer-to-peer electronic cash system. <https://bitcoin.org/bitcoin.pdf>
- Snapper, J. W. (1999): On the Web, plagiarism matters more than copyright piracy. In: *Ethics and Information Technology*, 1 (2), 127–135.
- Swan, M. (2015): *Blockchain: Blueprint for a New Economy*. O'Reilly Media, Inc.

Stereotype and Most-Popular Recommendations in the Digital Library *Sowiport*

*Joeran Beel*¹, *Siddharth Dinesh*², *Philipp Mayr*³,
*Zeljko Carevic*⁴, *Jain Raghvendra*⁵

¹ Trinity College Dublin, SCSS/KDEG, ADAPT Centre, Ireland
joeran.beel@adaptcentre.ie

² Birla Institute for Technology and Science, India
dinesh@mr-dlib.org

³ GESIS – Leibniz Institute for the Social Sciences, Germany
{philipp.mayr, zeljko.carevic}@gesis.org

⁴ National Institute of Informatics (NII), Japan
jain@nii.ac.jp

Abstract

Stereotype and *most-popular* recommendations are widely neglected in the research-paper recommender-system and digital-library community. In other domains such as movie recommendations and hotel search, however, these recommendation approaches have proven their effectiveness. We were interested to find out how stereotype and most-popular recommendations would perform in the scenario of a digital library. Therefore, we implemented the two approaches in the recommender system of *GESIS*' digital library *Sowiport*, in cooperation with the recommendations-as-a-service provider *Mr. DLib*. We measured the effectiveness of most-popular and stereotype recommendations with click-through rate (CTR) based on 28 million delivered recommendations. Most-popular recommendations achieved a CTR of 0.11%, and stereotype recommendations achieved a CTR of 0.124%. Compared to a “random recommendations” baseline (CTR 0.12%), and a content-based filtering baseline (CTR 0.145%), the results are discouraging. However, for reasons explained in the paper, we concluded that more research is

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 96–108.

necessary about the effectiveness of stereotype and most-popular recommendations in digital libraries.

Keywords: recommender systems; digital libraries; evaluation; stereotype recommendations; most-popular recommendations; content-based filtering

1 Introduction

Recommender systems for research papers typically apply content-based filtering, item-based collaborative filtering, co-occurrence calculations or graph-based recommendations (Beel et al., 2015a). Two less common recommendation classes are stereotyping and most-popular recommendations. Stereotyping is one of the earliest user modeling and recommendation classes. In a stereotype recommender system, some generalizing assumptions are made about users (e.g. males like cars and females like perfume), and then items are recommended that presumably are interesting for those stereotype users. A most-popular recommender system adopts a one-fits-all approach and recommends items that have the highest popularity. For instance, a news website could recommend those news articles that were most often read or that had the highest average rating over all users (Lommatzsch et al., 2016). The basic assumption behind such a recommender system is that users will like what most other users read, download, like, etc.

Both stereotype and most-popular recommendations received little attention in the community of research-paper recommender systems, although the two recommendation classes proved effective in other domains (Kay, 2000; Kobsa, 1993; Kobsa, 2001; Lamche et al., 2014; Mattioli, 2012; Rich, 1979). Our research goal is to explore the effectiveness of stereotype and most-popular recommendations in digital libraries, more specifically in GESIS' digital library *Sowiport*¹. The research question we attempt to answer is:

How effective are “Stereotype” and “Most-Popular” recommendations for recommending scholarly literature in digital libraries, Sowiport respectively?

¹ <http://sowiport.gesis.org>

2 Related work²

2.1 Stereotype recommendations

Stereotype recommendations were introduced by Rich in the book-recommender system *Grundy* (Rich, 1979). Rich was inspired by stereotypes from psychology where stereotypes allowed psychologists to quickly judge people based on a few characteristics. For instance, Rich assumed that male users have “a fairly high tolerance for violence and suffering, as well as a preference for thrill, suspense, fast plots, and a negative interest in romance”. Consequently, Grundy’s stereotype recommendation approach recommended action books and thrillers to male users.

One major problem with stereotypes is that they may pigeonhole users. While many men have a negative interest in romance, certainly not all do. In addition, building stereotypes is often labor intensive, as the items typically need to be manually classified for each stereotype. This limits the number of e.g. books that could be recommended (Barla, 2011).

Advocates of stereotypes argue that once the stereotypes are created, the recommender system needs little computing power and may perform quite well in practice. For instance, Weber and Castillo (2010) observed that female users were usually searching for the composer Richard Wagner when they entered the search query ‘Wagner’ on *Yahoo!*. In contrast, male users entering the same query usually were looking for the Wagner paint sprayer. Weber & Castillo modified *Yahoo!*’s search algorithm to show the Wikipedia page for Richard Wagner to female users, and the homepage of the Wagner paint sprayer company to male users searching for ‘Wagner’. As a result, user satisfaction increased. Similarly, the travel agency *Orbitz* observed that Macintosh users were “40% more likely to book a four- or five-star hotel than PC users” and when booking the same hotel, Macintosh users booked the more expensive rooms (Mattioli, 2012). Consequently, Orbitz assigned its website visitors to either the “Mac User” or “PC user” stereotype, and Mac users received recommendations for pricier hotels than PC users. All parties benefited – users received more relevant search results, and Orbitz received higher commissions.

² Some explanations of stereotype and most-popular recommendations are from Beel et al. (2015).

In the domain of research-paper recommender systems, stereotype recommendations have only been applied in the recommender system of the reference manager *Docear* (Beel et al., 2014; Beel et al., 2015b). The developers of the recommender system manually created a list of books and research articles relating to academic writing, and these documents were then recommended to the users of *Docear*. The authors report a mediocre effectiveness of the stereotype approach with an average click-through rate of 3.08%. In contrast, a standard content-based filtering approach achieved click-through rates slightly below 4%, and a novel content-based filtering approach, tailored to the users of *Docear*, achieved click-through rates around 7% (for more details about click-through rate as evaluation metric, please refer to Beel and Langer (2015) and the methodology section of the current paper).

We see a need for further research on stereotype recommendations in the domain of digital libraries. The *Docear* team recommended only documents about one topic, i.e. academic writing, and the recommendations were only tested in *Docear*. However, recommendation approaches may perform very differently in different scenarios (Beel et al., 2016; Beel et al., 2013). Therefore, we see the need to conduct the research in a different scenario than *Docear*, and with additional topics than academic writing.

2.2 Most-popular recommendations

In the domain of research-paper recommender systems, several recommender systems use popularity as an *additional* ranking factor (Bethard & Jurafsky, 2010; He et al. 2010; Ren, 2016; Totti et al., 2016; Zarrinkalam & Kahani, 2013). These systems first determine a list of recommendation candidates, for instance, with content-based filtering. Then, the recommendation candidates are re-ranked based on document popularity. For instance, out of the 20 recommendation candidates that are calculated with content-based filtering, the ten most cited papers might be recommended. Common metrics to calculate popularity are PageRank (Bethard & Jurafsky, 2010), HITS (He et al., 2010), Katz (ibid.), citation counts (Bethard & Jurafsky, 2010; He et al., 2010; Rokach et al., 2013), venues' citation counts (Bethard & Jurafsky, 2010; Rokach et al., 2013), citation counts of the authors' affiliations (Rokach et al., 2013), authors' citation count (Bethard & Jurafsky, 2010; Rokach et al., 2013), h-index (Bethard & Jurafsky, 2010), and recency of articles (ibid.).

To the best of our knowledge, there is no research on how effective it is to recommend items in a digital library only based on the items' popularity (e.g. loans, views, downloads, citations).

3 Methodology

For our research we used the digital library *Sowiport* (Hienert, Sawitzki & Mayr, 2015). *Sowiport* is operated by 'GESIS – Leibniz-Institute for the Social Sciences', which is the largest infrastructure institution for the Social Sciences in Germany. *Sowiport* contains about 9.6 million literature references and 50,000 research projects from 18 different databases, mostly relating to the social and political sciences. Literature references usually cover keywords, classifications, author(s) and journal or conference information and if available: citations, references and links to full texts. On a weekly base, *Sowiport* reaches around 22,000 unique users. These users spend on average 2 minutes in the system.

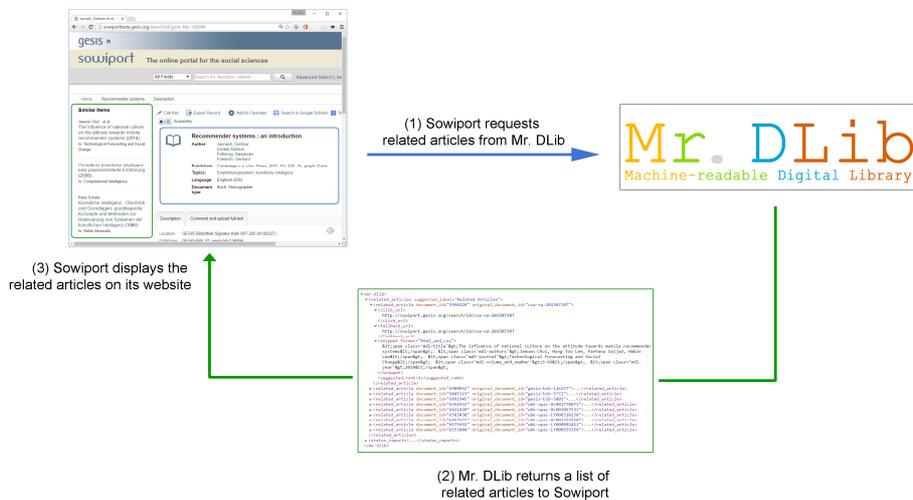


Fig. 1 The recommendation process of Sowiport and Mr. DLib

Sowiport co-operates with Mr. DLib³, an open Web Service to provide scholarly literature-recommendations-as-a-service (fig. 1). This means that

3 <http://mr-dlib.org>

all computations relating to the recommendations run on Mr. DLib's servers, while the presentation takes place on Sowiport's website.

Our recommender system shows related-article recommendations on each article's detail page in Sowiport (fig. 2). Whenever such a detail page is requested by a user, the recommender system randomly chooses one of four recommendation approaches to generate recommendations:⁴ 1. stereotype recommendations, 2. most popular recommendations, 3. content-based filtering (CBF), and 4. "random" recommendations, whereas CBF and random recommendations served as baselines. For content-based filtering recommendations, we used Lucene's "More Like This" function, a recommendation approach that is used by many research-paper recommender systems (Beel et al., 2015a). When the random approach is chosen, the recommender system randomly picks some documents out of the 9.6 million documents in the recommendation corpus.

The screenshot shows the Sowiport website interface. At the top, there is a search bar and navigation links. The main content area is divided into two columns. The left column, titled 'Similar Items', lists three books: 'Postmodernism, Marxism, and the Critique of Modern Economic Thought (1994)' by Jack Amariglio et al., 'Modernization and Postmodernization: Cultural, Economic, and Political Change in Forty-Three Societies (1998)' by Michael Hout, and 'Theoretical Orientations of Spanish Psychotherapists: Integration and Eclecticism as Modern and Postmodern Cultural Trends. (2006)' by Amparo Coscolla et al. The right column displays the details for the book 'MODERNIZATION AND POSTMODERNIZATION: CULTURAL, ECONOMIC, AND POLITICAL CHANGE IN 43 SOCIETIES' by Ronald Inglehart. Below the book details, there are buttons for 'Description', 'Cited by (275)', and 'Upload'. A callout box titled 'Related-Article Recommendations from Mr. DLib' is positioned above the 'Similar Items' section, with a red arrow pointing to it.

Fig. 2

Screenshot of Sowiport's website with recommendations in the left part of the page

⁴ The approaches are chosen with different probabilities. For instance, random-recommendations were only chosen with a probability of 4% because we needed these kinds of recommendations only as baseline.

To create stereotype recommendations, we assumed that a major part of Sowiport users – who are mostly students and researchers – are interested in the topics “academic writing”, “research methods”, and “peer review & research evaluation”. We used Sowiport’s search function to find 16 documents that we considered to be relevant for the three research topics, and these documents were then recommended to the users of Sowiport. Figure 3 shows more details about the 16 documents.

	Sowiport ID	Title	Year	Language
Academic Writing	dzi-solit-000215431	Erfolgreiches wissenschaftliches Schreiben	2015	de
	dzi-solit-0129221	Kreatives wissenschaftliches Schreiben: Tipps und Tricks gegen Schreibblockaden	2001	de
	fis-bildung-1018973	Writing for peer reviewed journals	2013	en
	fis-bildung-1068313	Kreatives Schreiben von Diplom- und Doktorarbeiten	1998	de
	fis-bildung-1071788	Kreatives wissenschaftliches Schreiben	2001	de
	fis-bildung-621436	Geniale Notizen	2002	de
Peer Review	gesis-bib-126169	Erfolgreiches wissenschaftliches Arbeiten: Seminararbeit, Bachelor-/Masterarbeit (Diplomarbeit), Doktorarbeit	2008	de
	csa-sa-201609258	Wissenschaftliches Publizieren: Peer Review	2014	de
	gesis-ssoar-2362	Exzellenz und Evaluationsstandards im internationalen Vergleich	2007	de
	gesis-ssoar-2530	Einleitung: Wie viel (In-)Transparenz ist notwendig? Peer Review Revisited	2006	de
Research Methods	gesis-ssoar-733	Peer Review in der DFG: die Fachkollegiaten	2007	de
	fis-bildung-949616	Empirische Forschungsmethoden	2010	de
	gesis-solis-00569924	Einführung in die Wissenschaftstheorie	2014	de
	gesis-solis-00598617	Forschungsmethoden und Statistik: ein Lehrbuch für Psychologen und Sozialwissenschaftler	2013	de
	gesis-solis-00606948	Forschungsmethoden	2013	de
	iab-litdok-K110511315	Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft	2010	de

Fig. 3 Details on the 16 documents that we selected as stereotype recommendations

	Sowiport ID	Title	Year	Language
Top Views	fis-bildung-999945	Guter Chemieunterricht	2013	de
	gesis-solis-00560882	Die Gesellschaft und ihre Gesundheit: 20 Jahre Public Health in Deutschland ; Bilanz und Ausblick einer Wissenschaft	2011	de
	gesis-solis-00551750	Thrillslider: Rutschen, Rausch und Rituale auf Spielplätzen, Festplätzen und in Aqua-Parks	2010	de
	gesis-solis-00526599	Weiterbildungsbeteiligung von Menschen mit Migrationshintergrund in Deutschland	2009	de
	fis-bildung-840181	Kommt der Herbst mit bunter Pracht	2008	de
Top Exported	...			
	gesis-solis-00605639	Organisieren am Konflikt: Tarifaueinandersetzungen und Mitgliederentwicklung im Dienstleistungssektor	2013	de
	gesis-solis-00606019	Soziale Arbeit und Stadtentwicklung: Forschungsperspektiven, Handlungsfelder, Herausforderungen	2013	de
	gesis-solis-00580567	Fokusgruppen in der empirischen Sozialwissenschaft: von der Konzeption bis zur Auswertung	2012	de
	gesis-solis-00563254	Handbuch zur Verwaltungsreform	2011	de
gesis-solis-00568965	Die Zukunft auf dem Tisch: Analysen, Trends und Perspektiven der Ernährung von morgen	2011	de	
...				

Fig. 4 Details on the most viewed and exported documents (excerpt)

For the most-popular recommendations we used two metrics to measure popularity. First, “views”, which measure how often a document’s detail-page was accessed by a visitor on Sowiport’s website. Second, “exports”, which measure how often documents’ metadata was exported on Sowiport’s website as e.g. BibTeX, EndNote, or email. For both metrics, we identified the 50 most popular documents for the month August, and recommended these documents to the users of Sowiport. Figure 4 shows some of the 2×50 documents, a complete list is available from us upon request.

We measured the effectiveness of the recommendation approaches with click-through rate (**CTR**). CTR describes the ratio of delivered to clicked recommendations. For instance, when 10,000 recommendations based on CBF were delivered, and 50 of these recommendations were clicked, the average CTR of CBF would be $50 \div 10,000 = 0.5\%$. The assumption is that the higher the CTR, the more effective is the recommendation approach. There is some discussion to what extent CTR is appropriate for measuring recommendation effectiveness, but overall it has been demonstrated to be a meaningful and well-suited metric (Beel & Langer, 2015; Joachims et al., 2005; Schwarzer et al., 2016).

Table 1: Number of displayed and clicked recommendations by recommendation approach

	Total	Content Based Filtering	Most Popular			Stereotype				Random
			Top Views	Top Exports	Overall	Academic Writing	Research Methods	Research Evaluation	Overall	
Displayed	28,214,883	24,335,531	1,187,845	1,060,647	2,248,492	149,235	147,034	84,938	381,207	1,249,653
Clicks	31,872	27,423	1,373	1,107	2,480	175	192	107	474	1,495

Between 17 October 2016 and 28 December 2016, Mr. DLib’s recommender system delivered 28,214,883 recommendations to Sowiport.⁵ Whenever comparing results of different algorithms, we report the significance level p , which is calculated with a two-tailed t-test. All data relating to this paper is available on *Harvard’s Dataverse*,⁶ including a list of the delivered and clicked recommendations as CSV file, the *R* script to analyze the data,

⁵ Whenever an article’s detail page was shown to a user, Mr. DLib returned between 1 and 15 related-article recommendations. Numbers include recommendations delivered to bots which crawled the Sowiport website. Clicks were recorded via JavaScript. Hence, click-through rates overall are rather low. Numbers include only recommendations which required 3 or fewer seconds to calculate because in the other cases we could not be sure that the recommendations were actually displayed to a user.

⁶ https://dataverse.harvard.edu/dataverse/Mr_DLib

and the figures and tables presented in this paper as PNG and CSV files (Beel et al., 2017).

4 Results

Figure 5 shows the click-through rates for the four recommendation approaches. Content-based filtering performed best with an average CTR of 0.145%, compared to a CTR of 0.12% for random recommendations ($p = 0.03$). Stereotype recommendations performed second best with a CTR of 0.124% on average, which is an improvement compared to random recommendations, however, with low significance ($p = 0.47$). Most-popular recommendations were even slightly less effective (CTR = 0.11%) than random recommendations, with high statistical significance ($p = 0.01$).

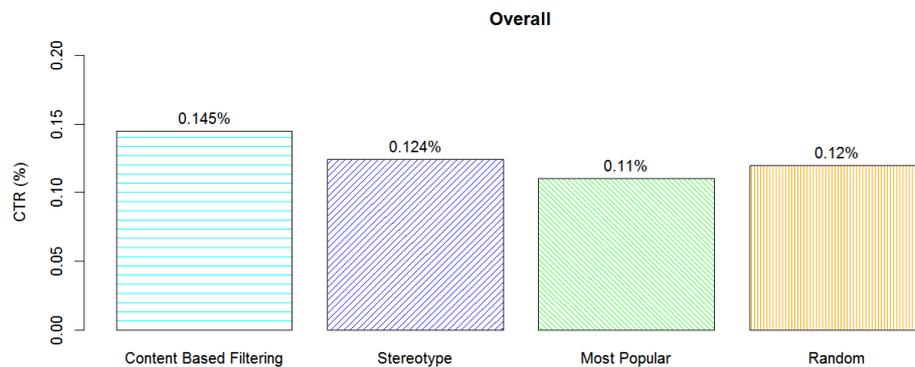


Fig. 5 CTR for the different recommendation approaches

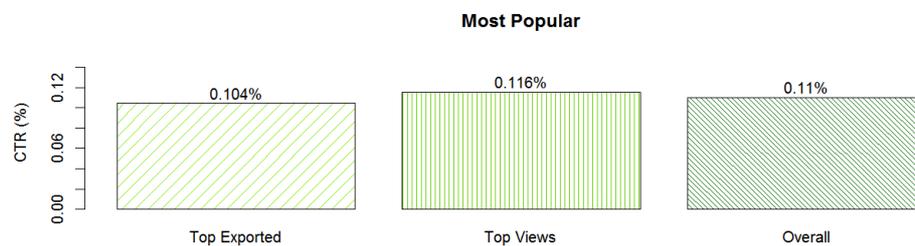


Fig. 6 CTR for the most-popular recommendation categories

For the most-popular recommendations it made no difference whether we used *exports* or *views* to determine the most popular recommendations (cf. fig. 6). CTR was 0.104% and 0.116% respectively, i.e. both CTRs are below CTR of random recommendations.

Looking at stereotype recommendations in detail reveals that CTR for the different categories varied (fig. 7). Recommendations for scholarly literature about academic writing achieved the lowest CTR (0.117%) among the stereotype recommendations. Recommendations about peer review and research evaluation achieved CTRs of 0.126%, and recommendations for literature about research methods achieved performed best with a CTR of 0.131%. However, the differences are statistically not significant.

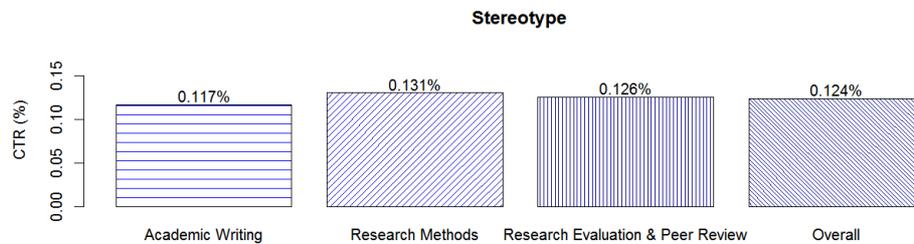


Fig. 7 CTR for the different categories of stereotype recommendations

5 Conclusion and outlook

Overall, the results are somewhat disappointing. Stereotype recommendations were about as (in)effective as random recommendations with both having a CTR of 0.124% and 0.12% respectively. This result contradicts previous research about stereotype recommendations from the Docear researchers. Most-popular recommendations were even statistical significantly less effective (CTR = 0.11%) than random recommendations.

Based on the current results, it seems not sensible to apply stereotype and most-popular recommendations, at least not on Sowiport. However, to reach a final conclusion we consider more research to be necessary. Among others, additional evaluation metrics might be sensible. In addition, a better detection of web spiders crawling the Sowiport website (and hence requesting recommendations), would lead to more reliable data. It might also make sense to experiment with other popularity metrics than views and exports and longer

or shorter periods of time to define a popular item. One interesting metric might be “libcitations” (White et al., 2009). Libcitations count a libraries’ stock of a given book and give an indicator of its popularity in that library. In addition, the effectiveness of most-popular recommendations could be researched in other scenarios, for instance in smaller libraries with a more homogenous user base.

Further research about stereotype recommendations could focus on identifying, which type of items (e.g. research articles, reviews, blog posts, news, software tools, or research projects) and which kind of topics researchers are most interested in. It could also be interesting to build more tailored stereotypes. Currently, we only had one ‘class’ of stereotypes, i.e. we assumed that all Sowiport visitors had the same interests in academic writing etc. If the recommender system knew, for instance, a visitor’s academic status (e.g. professor, post-doc, PhD student) or research discipline, the stereotype recommendations could be tailored better to the different user groups’ needs.

Acknowledgements

This work was supported by a fellowship within the postdoc-program of the German Academic Exchange Service (DAAD). This publication also has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under Grant Number 13/RC/2106. We are further grateful for the support provided by Sophie Siebert and Stefan Feyer.

References

- Barla, M. (2011): Towards social-based user modeling and personalization. In: *Information Sciences and Technologies – Bulletin of the ACM Slovakia*, 3, 52–60.
- Beel, J., and S. Langer. (2015): A Comparison of Offline Evaluations, Online Evaluations, and User Studies in the Context of Research-Paper Recommender Systems. In: *Proceedings of the 19th International Conference on Theory and Practice of Digital Libraries (TPDL)*, Lecture Notes in Computer Science (Vol. 9316) (pp. 153–168). [doi:10.1007/978-3-319-24592-8_12](https://doi.org/10.1007/978-3-319-24592-8_12)
- Beel, J., S. Langer, A. Nürnberger and M. Genzmehr (2013): The Impact of Demographics (Age and Gender) and Other User Characteristics on Evaluating Recommender Systems. In: *Proceedings of the 17th International Conference on*

- Theory and Practice of Digital Libraries (TPDL 2013)* (pp. 400–404). Valletta, Malta: Springer.
- Beel, J., S. Langer, B. Gipp, and A. Nürnberger (2014): The Architecture and Datasets of Docear’s Research Paper Recommender System. In: *D-Lib Magazine*, 20 (11/12). [doi:10.1045/november14-beel](https://doi.org/10.1045/november14-beel)
- Beel, J., B. Gipp, S. Langer, and C. Breitingner (2015a): Research Paper Recommender Systems: A Literature Survey. In: *International Journal on Digital Libraries*, 1–34. [doi:10.1007/s00799-015-0156-0](https://doi.org/10.1007/s00799-015-0156-0)
- Beel, J., S. Langer, G. M. Kapitsaki, C. Breitingner, and B. Gipp (2015b): Exploring the Potential of User Modeling based on Mind Maps. In: *Proceedings of the 23rd Conference on User Modelling, Adaptation and Personalization (UMAP)*, Lecture Notes of Computer Science (Vol. 9146) (pp. 3–17). Springer. [doi:10.1007/978-3-319-20267-9_1](https://doi.org/10.1007/978-3-319-20267-9_1)
- Beel, J., C. Breitingner, S. Langer, A. Lommatzsch, and B. Gipp (2016): Towards Reproducibility in Recommender-Systems Research. In: *User Modeling and User-Adapted Interaction (UMUAI)*, 26 (1), 69–101. [doi:10.1007/s11257-016-9174-x](https://doi.org/10.1007/s11257-016-9174-x)
- Beel, J., S. Dinesh, P. Mayr, Z. Carevic, and J. Raghvendra (2017): Stereotype and Most-Popular Recommendations in the Digital Library Sowiport [Dataset]. Harvard Dataverse. [doi:10.7910/DVN/HFIV1A](https://doi.org/10.7910/DVN/HFIV1A)
- Bethard, S., and D. Jurafsky (2010): Who should I cite: learning literature search models from citation behavior. In: *Proceedings of the 19th ACM international conference on Information and knowledge management* (pp. 609–618). ACM.
- He, Q., J. Pei, D. Kifer, P. Mitra, and L. Giles (2010): Context-aware citation recommendation. In: *Proceedings of the 19th international conference on World Wide Web* (pp. 421–430). ACM.
- Hienert, D., F. Sawitzki, and P. Mayr (2015): Digital Library Research in Action – Supporting Information Retrieval in Sowiport. In: *D-Lib Magazine*, 21 (3/4). [doi:10.1045/march2015-hienert](https://doi.org/10.1045/march2015-hienert)
- Joachims, T., L. Granka, B. Pan, H. Hembrooke, and G. Gay (2005): Accurately interpreting clickthrough data as implicit feedback. In: *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 154–161). ACM.
- Kay, J. (2000): Stereotypes, student models and scrutability. In: *International Conference on Intelligent Tutoring Systems* (pp. 19–30). Springer.
- Kobsa, A. (1993): User modeling: Recent work, prospects and hazards. In: *Human Factors in Information Technology*, 10, 111–111.
- Kobsa, A. (2001): Generic user modeling systems. In: *User modeling and user-adapted interaction*, 11 (1–2), 49–63.

- Lamche, B., E. Pollok, W. Wörndl, and G. Groh (2014): Evaluating the Effectiveness of Stereotype User Models for Recommendations on Mobile Devices. In: *UMAP Workshops*.
- Lommatzsch, A., N. Johannes, J. Meiners, L. Helmers, and J. Domann (2016): Recommender Ensembles for News Articles based on Most-Popular Strategies. In: *Working Notes of the 7th International Conference of the CLEF Initiative*, Evora, Portugal.
- Mattioli, D. (2012): On Orbitz, Mac users steered to pricier hotels. In: *Wall Street Journal*. <http://online.wsj.com/news/articles/SB10001424052702304458604577488822667325882>
- Ren, X. (2016): Effective citation recommendation by information network-based clustering.
- Rich, E. (1979): User modeling via stereotypes. In: *Cognitive science*, 3 (4), 329–354.
- Rokach, L., P. Mitra, S. Kataria, W. Huang, W. and L. Giles (2013): A Supervised Learning Method for Context-Aware Citation Recommendation in a Large Corpus. In: *Proceedings of the Large-Scale and Distributed Systems for Information Retrieval Workshop (LSDS-IR)* (pp. 17–22).
- Schwarzer, M., M. Schubotz, N. Meuschke, C. Breitingner, V. Markl, and B. Gipp (2016): Evaluating Link-based Recommendations for Wikipedia. In: *Proceedings of the 16th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL), JCDL '16* (pp. 191–200). Newark, N.J.: ACM. [doi:10.1145/2910896.2910908](https://doi.org/10.1145/2910896.2910908)
- Totti, L. C., P. Mitra, M. Ouzzani, and M.J. Zaki (2016): A Query-oriented Approach for Relevance in Citation Networks. In: *Proceedings of the 25th International Conference Companion on World Wide Web* (pp. 401–406). International World Wide Web Conferences Steering Committee.
- Weber, I., and C. Castillo (2010): The demographics of web search. In: *Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval* (pp. 523–530). ACM.
- White, H. D., S. K. Boell, H. Yu, M. Davis, C. S. Wilson, and F. T. H. Cole (2009): Libcitations: A measure for comparative assessment of book publications in the humanities and social sciences. In: *Journal of the American Society for Information Science and Technology*, 60 (6), 1083–1096. [doi:10.1002/asi.21045](https://doi.org/10.1002/asi.21045)
- Zarrinkalam, F., and M. Kahani (2013): SemCiR – A citation recommendation system based on a novel semantic distance measure. In: *Program: electronic library and information systems*, 47 (1), 92–112.

Session 3:

User Perceptions of Information Systems

Content, Physical Appearance, Copy Condition

Tagging Customer Book Reviews

Tjaša Jug, Maja Žumer

University of Ljubljana, Faculty of Arts, Slovenia
{Tjasa.Jug, Maja.Zumer}@ff.uni-lj.si

Abstract

Users of online bookstores are not interested only in general book description when searching and buying books, but also in subjective reader opinion, which could be found in online reviews. Reviewers usually comment on book content, but may also mention other aspects of the received book, such as binding, illustrations, translation etc. Meanwhile buyers are not always interested in the same aspect of a book, especially when they need it for a special purpose. Currently, obtaining non-content book information from reviews is difficult, therefore it would be reasonable to rethink their presentation and organization. In our study, we used an interview and a task solving method to determine whether social tagging could be an appropriate aid for this purpose. The results show that free tagging offers insight into users' vocabulary but is not optimal for online review presentation. Nevertheless, it represents a good basis for creation of categories that describe books on different levels of abstraction and could be used as a filtering tool, which would select only those reviews containing the aspects of a book a buyer is interested in.

Keywords: books; reviews; tagging; Amazon; book aspects; FRBR

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 110–115.

1 Theoretical framework

Customer reviews are often integrated in online bookstores and other book-related platforms as they represent an important factor for book purchasing and borrowing decision (Chevalier & Mayzlin, 2006; Lin, Huang & Yang, 2007; Huang & Yang, 2010). Some online bookstores, such as Amazon, have made a step further by combining all formats and editions of a work under a single title and all customer reviews for a work. This method somewhat resembles the Functional Requirements for Bibliographic Records (FRBR) (Functional ..., 1998) which considers a book at various levels of abstraction (*Work, Expression, Manifestation* and *Item*). Book reviews do not focus only on the *Work*-level aspects of a book, such as content, genre, author etc., but also on the characteristics of a particular edition. What is more, they can also comment on the received copy (an *Item*) or their buying experience (Jug & Žumer, 2016). Meanwhile, book buyers might be interested in various aspects of a book not only when they are searching for a book for a specific purpose, such as school reading assignment or a gift, but also when they can choose among versions with different translators, illustrations, text design or binding. Therefore, merging customer reviews may cause problems when reviewers express their opinion on aspects of a book that do not correspond to all versions of the book. Some researchers (Liu, Karahanna & Watson, 2011) have already found that it would be reasonable to group customer reviews into categories according to product attributes described in the reviews. Due customer interests in various book attributes, also online bookstores could use a filtering tool that would help individuals to select only comments describing a particular aspect of a book. As users know best what they need, it might be useful if they could classify reviews with social tagging and create categories for different book characteristics. We are assuming that user generated content could be helpful for this manner as it reflects their interests and vocabulary (Spiteri, 2007).

2 Research questions and methods

In our study, we aimed to identify which aspects of a book people recognize while reading book reviews and how do this aspects overlap with the FRBR

entities. The FRBR model was chosen because it was proven to be intuitive (Žumer, Salaba & Zhang, 2012) and it provides an appropriate clustering mechanism. We were also interested in vocabulary people use to describe different book characteristics recognized in the comments and whether a social tagging would be an appropriate method for different review organization.

Prior to the study, we modified 11 Amazon reviews of 9 well known book titles with many versions in a way that they contain opinions on diverse book aspects and include different combinations of book attributes. The most of the reviews commented on printed books, but there were also two comments for audio books and one for a kindle version. After preparing the reviews, we carried out a content analysis, where we determined which FRBR entities match attributes mentioned in our comments. In the second step of our study, twenty-five people read reviews and added their own tags to describe the book attributes they had identified in the reviews. After the data collection, we compared obtained tags to the FRBR entities and designed 10 categories in a way that covers all four levels of abstraction of the FRBR model and at the same time reflects participants' terminology. Names of the categories were based on the most frequent words assigned by participants to the part of the text describing the book on the particular level of abstraction. For example, category *Content* applies to FRBR's *Work*, *Edition* to *Expression* or *Manifestation*, *Medium*, *physical appearance* and *Additions* to *Manifestation* and *Copy condition* or *Shopping experience* to *Item*. We also added category *Target audience* and categories *Comparison with different editions* and *Comparison with different works*, which describe related products.

One month after the first round, we asked the same participants to read the same reviews and to identify book aspects mentioned in them by marking relevant categories from the list. We also tested the applicability of these categories with the same set of reviews and a new group of twenty-five participants in round three and compared the detected aspects in three different rounds. During the study we did not offer the participants any other information about the categories, as we wanted to test how self-explanatory are they.

Figure 1 represents a review for the special edition of *The Hobbit*. In this review, we identified FRBR entity *Work* where the reviewer mentions the book's content. While he likes the cover, binding and illustrations, which are attributes of a *Manifestation*, his disappointment and rating refer to the received copy – an *Item*. He also mentions *Related work* (*Lord of the Rings*).

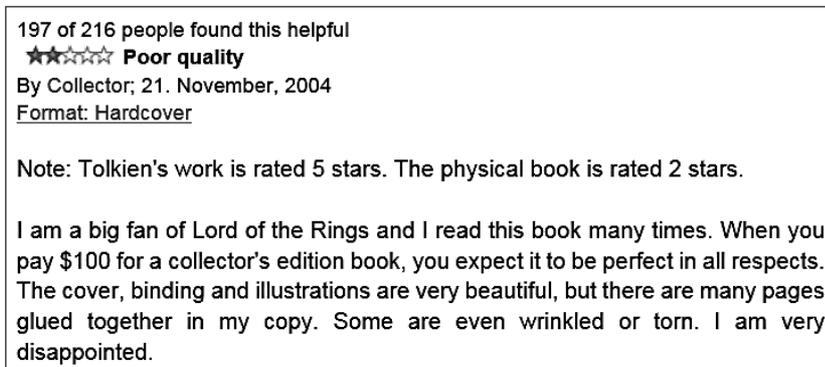


Fig. 1 Example of a review used in our study

In figure 2 one can see some of the tags and categories assigned to the review from figure 1 by participants in two rounds. For this review, first group of the participants assigned 78 unique tags in round 1. As we can see in the middle column, we obtained many tags with similar meanings and low frequencies, such as *Collector's edition*, *Special edition*, *Physical edition* etc., which would not be appropriate for automatic processing of book reviews. As seen in the right column, choosing from a predefined list in round 2 proved to be a better approach as these categories are easy to select and are therefore more transparently reflecting the most noticeable and important aspects of a book in a specific review.

FRBR	FREE TAGS	CATEGORIES
Work (15)	Content (4) Work evaluation (2)	
Manifestation (16)	Collector's edition (4) Physical appearance (2) Special edition (1) Physical edition (1) Concrete edition (1)	Edition (14) Physical appearance (18)
Item (23)	Glued pages (3) Evaluation (2) Condition (1) Product quality (1) Bad shopping experience (1)	Copy condition (22) Shopping experience (16)
Related work (3)	Other work (1) Continued by (1)	
	Personal opinion (3)	

Fig. 2 FRBR entities, tags and categories added in both rounds to the review from figure 1 (number of occurrences)

3 Research results

After each round, we asked the participants to determine the difficulty of the task and whether they would use the tagging system if it were integrated in the online bookstore. Figure 3 shows their responses. The most participants from the first round consider free tagging as moderate or difficult (22) and they said that they would not tag customer reviews in this way (16). On the contrary, they said the task was easy (14) when they had the list of categories and that they would use this tool (16), especially for the books they are interested in. The results were similar in the third group where the participants did not use free tagging method.

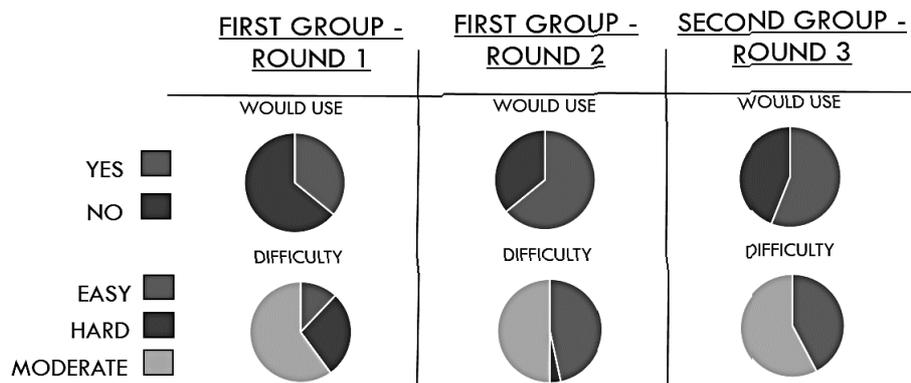


Fig. 3 Use of tags and difficulty of the task

The participants were also asked to comment on the pre-made list of the categories. Most were satisfied with the number of categories, but they had problems understanding some of them. They were not sure which book information falls under the term *Additions*. They also said that they are unsure about categories *Target audience* and *Medium*, although the results show that they chose these categories properly. Some participants also did not distinguish between categories *Physical appearance* and *Copy condition* and therefore often marked them both. They also suggested some new categories, for example *Subjective opinion*, *Value*, *Comparison with different medium*, *Technical details* and a *Free category*, where they could add their own tag.

4 Conclusion

The results indicate that people consider book information at various abstraction levels that match those expressed in the FRBR model. We also found that free tagging could be an effective method for identification of book attributes mentioned in the reviews but is not optimal for their organization and presentation. Meanwhile, a controlled list of categories that represent different aspects of a book could be a useful tool for this purpose and would enable automatic grouping and filtering of reviews. However, we cannot be sure if this technique would actually be used in an online bookstore context, therefore more research on a bigger sample and on an actual implementation of this tool would be needed.

References

- Chevalier, J. A. and D. Mayzlin (2006): The effect of word of mouth on sales: Online book reviews. In: *Journal of marketing research*, 43 (3), 345–354.
- Functional Requirements for Bibliographic Records: Final report* (2009). Munich: K.G. Saur Verlag.
- Huang, Y. and W. I. Yang (2010): A study of internet book reviews and borrowing intention. In: *Library Review*, 59 (7), 512–521.
- Jug, T. and M. Žumer (2016): Do We Need Better Online Book Review Organisation? In: *Libraries in the digital age (LIDA)*, Zadar.
- Lin, T. M., Y. K. Huang and W. I. Yang (2007): An experimental design approach to investigating the relationship between Internet book reviews and purchase intention. In: *Library & Information Science Research*, 29 (3), 397–415.
- Liu, Q. B., E. Karahanna, and R. T. Watson (2011): Unveiling user-generated content: Designing websites to best present customer reviews. In: *Business Horizons*, 54 (3), 231–240.
- Spiteri, L. F. (2007): The structure and form of folksonomy tags: The road to the public library catalog. In: *Information technology and libraries*, 26 (3).
- Žumer, M., A. Salaba, Y. Zhang (2012): User Verification of the FRBR Conceptual Model and Testing of FRBR Prototypes. In: *International Conference on Asian Digital Libraries* (pp. 148–151). Berlin: Springer.

Virtual Assistants

A Study on the Usability and User Perception of Customer Service Systems for E-Commerce

Valerie Claessen, Adrian Schmidt, Tamara Heck

Heinrich-Heine-University Düsseldorf, Germany
{valerie.claessen, adrian.schmidt, tamara.heck}@hhu.de

Abstract

Virtual assistants are able to support users in finding the right information. These programs use natural language processing, learning techniques and social abilities to offer adequate usability experiences for users. In e-commerce, virtual assistants are applied to support users in finding appropriate service information or products. This work evaluates the information service quality of three virtual assistants on e-commerce websites. The analyzed aspects cover service quality as well as user perception of virtual assistant systems. First results show that the apparent technology used in the construction of the virtual assistant has a substantial influence on user experience, by the users' perceived interaction with the assistant becoming more intuitive and therefore more enjoyable. Overall, all assistants were met with a general sense of enthusiasm. However, scores on the usefulness of the services show that they need to be improved regarding several relevant features.

Keywords: virtual assistant; intelligent agent; user study; human computer interaction; human information behavior

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 116–130.

1 Introduction

Human to human communication is progressively being replaced by human to computer interaction, while the prevalence of online sales and distribution is increasing rapidly. Advances in machine learning and data mining offer new opportunities in e-commerce to support and assist customers. One ongoing trend is the application of virtual assistants. Market research Gartner prognoses immense developments in intelligent apps, which assist companies and their employees (Panetta, 2016). In e-commerce, virtual assistants shall make users' tasks easier as well. Such programs are a further development of search engines (Chao et al., 2016) that have specific features like autonomy, proactiveness and learnability, and "assist and replace users in executing the time-consuming task of compiling information" (ibid.: 118). Those agents use natural language processing to communicate with users. In research literature, different terms are used to describe those agents: Virtual or intelligent agents, assistant or avatars, bots or even virtual personal agents, which might refer to software or real human assistants (compare McGoldrick, Keeling & Beatty 2008 for an overview). Modern programs have a further relevant feature, i.e. social abilities: An "intelligent agent interface [that] imitates several dimensions of social interaction" (Kuligowska & Lasek, 2005: 4). This software is socially interactive and simulates a kind of human to human dialogue. Such assistants are applied in the fields of education and e-learning, e-commerce, administration, and consulting. They shall advice users and/or lead them to buy products or services. In the following work, we use the term virtual assistant (VA) and concentrate on software that uses any natural language processing and has a graphical human-like, i.e. personified character. The evaluated VAs are current examples of assistants by three leading German companies active in the VA software development.¹ They are all inserted on e-commerce websites and support users in finding products or service information. Research focuses on the performance of these VAs and their user perception and usefulness.

¹ For an overview of VAs and companies see: <https://www.chatbots.org/country/de/>.

1.1 Related work

Studies on virtual assistants focus on diverse research aspects, like natural language processing and recommender system techniques (Chang, Lee & Wang, 2016; García-Serrano, Martínez & Hernández, 2004), VA interface and its user perception (Von der Pütten et al., 2010) and VA implementation for specific purposes like in e-learning or healthcare services (Ahamed et al., 2006). Compared to “physically-embodied agents” (also called robots), virtual agents are purely digital (Li, 2015). However, many research papers only refer to virtual agents or VAs as the software, i.e. the part of a system being able to process natural language and learning techniques from either text or spoken language. Other popular terms for such technical programs are “chatbot” or “chatterbot”. Besides this definition, other researcher assign a morphology to VAs, thus VAs can be either anthropomorphic, zoomorphic, caricatured or functional (Fong, Nourbakhsh & Dautenhahn, 2003). Furthermore, those VAs can have different types of modalities depending on their morphology: For example, an anthropomorphic VA (also called personified agent) can have an entire body, or just arms and a head (Li, 2015). VAs are distinguished from avatars, which represent and reflect the behavior of a human being, like avatars in Second Life (Von der Pütten et al., 2010). In this paper, we refer to VAs, which are represented by a digital anthropomorphic character. Besides research on VA techniques, studies assessing anthropomorphic VAs concentrate on users’ perception of these personified agent characters. In their user study, Shiban et al. (2015) tested the influence of two human-like agents (male and female) on e-learners’ enjoyment and interest. They detected no great differences between the agents, and students had similar performance even without the presence of any agent. Etemad-Sajadi (2014) studied ten diverse aspects (e.g. trust, ease of use, aesthetic) based on an expanded technology acceptance model (TAM) of a personified VA implemented on a restaurant’s website. The author’s analysis and his deployment of a user survey had similar intentions as the following study, with a focus on the usefulness and perception of a VAs. Results showed that usefulness and enjoyment positively influenced the users’ likelihood of revisiting the restaurant’s website. Furthermore, the aesthetic aspect of a VA and a user’s enjoyment have a strong interrelation, as well as trust and usefulness. Other studies, which implement and evaluate VAs, also have a focus on e-commerce and marketing aspects (e.g. Keeling, McGoldrick & Beatty, 2010). Their research questions mostly focus on the impact of VAs on the

quality of websites. The proposed work will have a deeper focus on human to VA interaction and a user's information need, leading to questions such as 'how do users evaluate information by a VA and how is information perceived via a VA?'. The study compares the usability of diverse VAs, regarding not only language processing techniques, but VA's performance and its perception from a user perspective. It will give first insights into human to VA interaction and its influence in user information behavior. The first results introduced in this paper concentrate on the perceived quality of the system and perceived quality of the content.

1.2 Research question and VAs

The following work addresses the VAs' performance and perception, and answers the following research question: How do VAs perform according to relevant information service criteria?

To answer this question, we evaluated and compared examples of three VAs, based on an empirical user study as well as a heuristic analysis (Sarodnick & Brau, 2011). We searched for companies offering German-speaking VA software technology and chose current and similar VAs deployed on German websites from three leading providers²: Carla³ by Kauz Linguistic Technologies, Clara by Novomind⁴, and Jana⁵ by Artificial Solutions⁶.

Carla has been active since early 2016 and was designed to be a fictional worker of a chocolate manufacturer. With Carla, customers are able to be recommended chocolate according to their preferences. The drawn female VA answers all kinds of questions regarding the creation of the offered goods, the manufacturer, as well as shipping and billing. The 3D rendered female VA of Otto, Clara⁷, went online in May 2013. Similar to Carla, Clara⁸ was designed to inform users of products and services of the Otto GmbH. Lastly, we considered the VA of the E-Post, Jana, which was deployed in December 2013. Jana is represented by a picture of a real woman. She has

2 For an overview see <https://www.chatbots.org/country/de>.

3 www.kauz.net/Carla

4 www.novomind.com

5 www.epost.de/privatkunden/hilfe-jana.html

6 www.artificial-solutions.com

7 www.otto.de/kontakt

8 To avoid any confusion with the VA Carla, Clara is called Otto from now on.

the role of an E-Post consultant and answers questions regarding the delivery and receiving of electronic mail by the Deutsche Post DHL Group.

2 Methods

The evaluation of the three VAs is based on the information service evaluation (ISE) model (Schumann & Stock, 2014) because it allows for a holistic view of diverse aspects of a service. ISE combines several traditions of evaluation and technology acceptance research and consists of five dimensions. In this study, we concentrate on dimension one, i.e. quality of information service, which includes perceived service as well as content quality evaluation and considers objective aspects (efficiency, effectiveness, gamification, functionality and usability) as well as users' perception (e.g. perceived usefulness and trust).

2.1 Heuristic usability evaluation

We applied a heuristic usability evaluation (Sarodnick & Brau, 2011: 144) to detect a VA's weaknesses relevant to the interaction with users. We determined ten relevant criteria covering the four objective categories in the ISE model (cf. table 1). We did not consider the aspect of gamification, as no VA offered any form thereof. After having contacted all three leading German developers, we were able to consult with Kauz Linguistic Technologies to ensure the design of a fitting criteria catalog, which could fit to the ISE model. To test the criteria, we designed queries corresponding to a VA's individual remit, i.e. VAs that specialize in sales were inquired about product consultation, VAs designed for information got questions regarding service. We determined the correctness of each reply given by the VA after having asked one query, with the exception for the categories "detail specification", "response calculation" and "context awareness". As the first two criteria require the VAs to interpret a combination of user information, two practice queries were asked first and a subsequent third one later, so that its reply by the VA could be evaluated. In terms of "context awareness", a dialogue with the VAs was established and their replies to the last query in that contextual dialogue was assessed.

For a service it is important that it is capable of “doing things right” (Schumann & Stock, 2014) to satisfy the customer or user. Schumann and Stock summarize relevant criteria in the category “efficiency”. Aspects like correctness, swiftness of the disclosure and quality of information are relevant for an information service efficiency. The VAs were tested for these qualities, especially in addition to the referral to possible useful and fitting sub-sites or pop-ups and the accurate usage of auto-correction.

Traditional effectiveness in retrieval research entails recall and precision studies (ibid.) – notably the last value considers a user’s subjective view. In our user study (see chapter 2.2), we directly asked participants, if the VAs gave sufficient answers. The heuristic part considers more objective aspects important for successful VAs, which are autonomous information fishing (the VA’s prompting for more and specific information from the user) and the interpretation of user intention (the VA’s independent, flexible and accurate prediction and understanding of the user’s input and needs).

As an additional aspect, Schumann and Stock depict the rating of the functionality as an “extent of its functions for information production and information searching”, which make a service more valuable. Thus, we tested the VAs for further features, such as the recognition and processing of negation and response calculation. Response calculation means the VA’s ability to combine different user input (e.g. the desired item, the cost of said item and the delivery options for that item) and provide appropriately accumulated information to the user, e.g.: User: “I’m looking for chocolate with fruits”, VA: “Then I can offer you the Choconegro 1001 Nights, the Choconegro Creme Cointreau and the Chocoblanco Raspberry Brittle”, user: “I prefer white chocolate”, VA: “I’d like to offer you the Chocoblanco Raspberry Brittle”, user: “I would like to buy white chocolate with fruits”, VA: “Then I’d suggest you try the Chocoblanco Raspberry Brittle. It’s our best white chocolate with fruits!”. In addition, the VAs were tested for context awareness, meaning that they can associate referrals to preceding inputs, e.g.: “I like how the white chocolate looks.” / “How much is it?” Here, “it” is referring to the white chocolate and the VA would still receive this input contextually and give a correct response.

Usability, i.e. serviceability of a VA, is the most important aspect for our study. In an ideal situation, the VA should be able to counteract any possible frustration on the part of its users. Therefore, the VA should identify a user’s sentiment, allow for flexible wording of any input, recognize details in a user’s inquiry and be aware of the context of each sentence or question.

We applied a scoring system to this evaluation. The highest achievable score for each criterion was set at 100 per cent, while the lowest most possible score resulted in 0 per cent. For each category we evaluated 50 VA replies to our queries. We considered two cases:

1. *Binary evaluation*: The examination of whether the reply of a VA was deemed correct in correspondence to the user input. It was only distinguished between ‘yes’ or ‘no’, or ‘correct’ and ‘incorrect’.
2. *Gradation evaluation*: In some cases, several replies given by the VA contained correct and incorrect part. In this case, scales with a range of 0 (incorrect) to 2 (correct) were implemented, where an almost correct answer could be considered with the score of 1.

Table 1: Assessment criteria and examples corresponding to the ISE categories

No.	Criteria	Examples	ISE-Category
1.	Auto Correction	‘What options do I have?’ → Understanding of the query despite the typo.	Efficiency
2.	Forwarding	Being forwarded to the developers when specifically asked to be.	Efficiency
3.	Autonomous Information Fishing	‘I need information.’ → The VA proactively asks the user what they need information on.	Effectiveness
4.	Interpretation of User Intention	‘I can’t find any contact information.’ → The VA provides the user with contact information instead of simply apologizing	Effectiveness
5.	Sentiment Detection	‘I like you.’ → The VA reacts positively.	Functionality
6.	Negation	‘Do you offer shirts without a v-neck?’ → The VA limits her offer to shirts without v-necks.	Functionality
7.	Response Calculation	‘Do you offer vegan chocolate without nuts?’ → The VA recognizes that the offered chocolate has to be vegan and without nuts at the same time.	Usability
8.	Detail Specification	‘How long does shipping take within Germany?’ → The VA delivers information on shipping, specifically in Germany.	Usability
9.	Context Awareness	User: ‘Do you offer a subscription?’ VA: ‘Yes we do.’ User: ‘How much would it cost?’ → The VA recognizes that ‘it’ refers to ‘subscription’.	Usability
10.	Flexible Wording	‘When was your company founded’ and ‘Since when is your company active?’ → The VA answers questions about a topic regardless of wording.	Usability

2.2 Empirical user study

The user study took place in March and May 2016 and included search scenarios followed by a survey. Each participant evaluated one VA and had to pass two tasks (limited to eight minutes each). First, the participants had time to make themselves comfortable with the website and the VA. Afterwards, they were given a defined task (performing a sales conversation). In the second task, the participants were requested to ask the VA questions regarding service, products and the company the VA represented. After each task, the participants answered survey questions corresponding to the task. The survey included 50 questions (single and multiple choice, rating scales, free text) and was divided into six parts: Demographics, appraisal and expectation of VAs, experience during task one and two regarding information quality, general experience after tasks, and preferences on VA aesthetic. Additionally, the participants' voices and the mouse movements on the screen were recorded. In the following section, we concentrate on results regarding the four system quality aspects of the ISE model, which lead to the following survey questions using a 7-point Likert scale:

1. Simplicity/ease of use: The interaction with the virtual assistant felt ... forced (1) – intuitive (7)
- 2a. Usefulness (questioned after 1st task): For my task, the information given by the VA were... useless (1) – useful (7)
- 2b. Usefulness (questioned after 2nd task): The information on products, services and the company was...useless (1) – useful (7)
3. Trust: I felt I was being taken seriously by the virtual assistant. I disagree (1) – I agree (7)
4. Fun: I had fun interacting with the virtual assistant. I disagree (1) – I agree (7)

3 Results

Table 2 shows the ten assessed criteria of the heuristic evaluation and the corresponding rate of correct uses by each VA in percentage. It has to be mentioned that the feature of forwarding to possibly helpful sub-sites was difficult to test because the quality of the sub-sites could not be assessed ap-

appropriately regarding the question the VAs were tested on. However, all VAs are able to link to any sub-sites.

Table 2: Assessed criteria of the heuristic evaluation

Category	Scores for Carla in %	Scores for Otto in %	Scores for Jana in %
1. Auto Correction	47.5	80	87.5
2. Forwarding	100	100	100
3. Autonomous Information Fishing	69.2	28.8	54
4. Interpretation of User Intention	68.2	45.5	63.6
5. Sentiment Detection	77.8	83.3	66.7
6. Flexible Wording	62.1	55	53.8
7. Response Calculation	100	25	25
8. Negation	90	20	30
9. Detail Specification	84.5	32.8	39.7
10. Context Awareness	100	33.3	66.7

Regarding the user study, 44 students (23 female, 21 male, age 17 to 34) from diverse faculties (whereof 31 users had information science as part of their studies) participated. Carla and Jana were evaluated by 15, Otto by 14 participants. 9 of the 44 persons stated that they had experience with VAs, half of all participants said that the usage of a VA must be fun. This statement assumes that enjoyment is indeed a relevant influencing aspect, as Etemad-Sajadi (2014) already found out. After the interaction with the VAs, the participants were not quite satisfied: The medians of Otto and Jana is 3, only Carla has a median of 5 (fun, fig. 1).

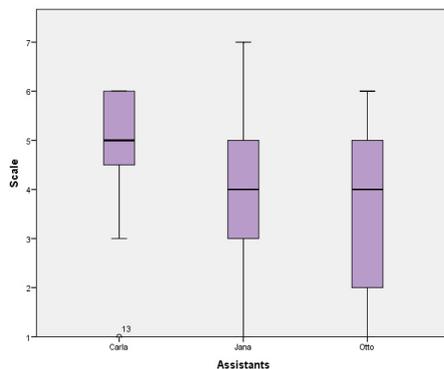


Fig. 1 Boxplots of the survey questions on simplicity, trust and fun

The aspects of ease of use and trust are also shown in figure 1, which shows a boxplot comparison of the data recorded through the 7-point scales. Simplicity shows a higher scattering of Carla's data yet with a direction towards higher scores. Carla's median of 4 only slightly contrasts with the median of 3 from the Jana and 3.5 from Otto. Regarding trust, Carla's median of 5 contrasts with Jana's and Otto's median of 4.

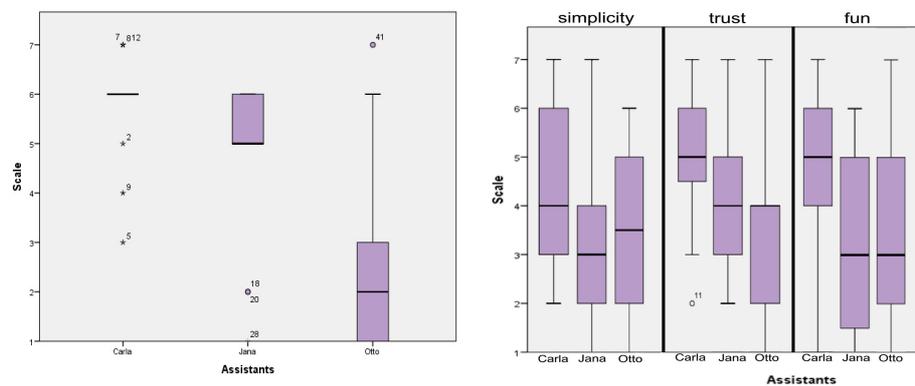


Fig. 2 Usefulness after the 1st task (left) and after the 2nd task (right)

Concerning the perceived usefulness after task one (performing a sales conversation; fig. 2 left), Carla's delivered information was rated to be higher than the other two VAs, with a median of 6. However, participants' assessment thereof was less unanimous than for Jana or Otto. While Jana was rated similarly to Carla, with a median of 5, Otto's was rated worst with a median of 2. Regarding perceived usefulness after task two (open questions), the right diagram (fig. 2) displays a higher dispersion in the recorded data concerning the usefulness of Otto's information. Participants perceived both Jana and Otto as generally less useful as indicated by their medians of 4 in comparison to the information delivered by Carla, with a median of 5.

4 Discussion

Judging by Carla's replies to our queries, it is assumed that her superior natural language understanding is due to her linguistic capabilities that were designed and implemented differently from the other VAs'. Parts of the VAs'

processing techniques could be derived from questioning the VAs based on the ten criteria. It became obvious that Carla's strongest abilities were being able to identify the combination of information and keeping a conversational flow by seemingly being aware of what has already been mentioned in the interaction with the user. Additionally, Carla demonstrated her ability to deliver user tailored information and her understanding of negation in inquiries. While other aspects were deemed as good, Carla's difficulty with fittingly auto correct user questions became obvious. Otto appeared very capable of understanding user moods and looking past typos and other mistakes, yet often failed to understand a negation and the combination of information in a user's inquiry. Jana demonstrated her ability to autocorrect typos very efficiently and detects indirectly articulated intentions by the user. Yet she appeared to have difficulties with understanding negations and combined information. More specifically, regarding the criteria of response calculation, Carla analyzed every word in its grammatical form and its relation to other words in the sentence, meaning that she was programmed to process and interpret the morphological and syntactical significance of the user input. This resulted in a more precise handling of the user request. While some predetermined replies were observed, many answers seem to be given in real time. In contrast, Otto and Jana accommodate the ability to filter out keywords from the user input and match them with pre-existing articles that either functioned as the answer to a user inquiry or that were attached as an extra text section in addition to a generic message. It became apparent that due to the lack of syntactic analysis, recognizable keywords would still be processed, regardless if the user prepends or appends unintelligible strings of characters and digits. If two or more triggering keywords were entered, Otto and Jana addressed only one. The order in which these keywords were entered did not appear to affect which keyword was processed by the two VAs. It remains uncertain, whether an internal priority list for all known keywords exists or whether Otto and Jana favor specific keywords. The feature of auto correction was present in all VAs, while it appeared to have functioned more reliably with Otto and Jana, most likely due to their keyword-only analysis, since the lack of additional linguistic processing had no interfering effect on this feature. With autonomous information fishing, Carla proved to be the strongest with Otto being the weakest, as it rarely prompted the user for more information. Regarding interpretation of user intention, the three VAs seemingly featured a rather comparable understanding of the purpose of inquiries. Concerning sentiment detection, Otto proved to be the most aware of a user's

emotions, while Jana remained the most unaware. Furthermore, Carla proved the most flexible with wording of inquiries, although all three VAs allowed for an average flexibility. It was also obvious that Carla understands negation (e.g. “Do you offer vegan chocolate without nuts?”), detail specifications (e.g. “How long does shipping take within Germany?”) and conversational context (e.g. anaphora resolution) the best, while Otto and Jana scored rather poorly in these categories.

The user study shows a similar trend concerning Carla’s favorable position. Regarding the aspect of simplicity, figure 2 shows a higher scattering of Carla’s data, yet with a direction towards higher scores, especially in comparison to Jana. The results of the participants’ judgment of how seriously they were taken (trust), suggest Carla’s distinct obligingness. Carla’s median of five contrasts with Jana and Otto’s median of four. Looking at the aspect of fun, interaction with Carla was deemed to be the most fun. In contrast, participants had slightly less fun with Otto and Jana, but their opinions were not unanimous, as the dispersion shows. Carla prevails in the assessment of usefulness, with Jana taking a close second place. Both VAs offer a wide palette of information on their offered services and goods. Otto’s significantly worse score is most likely attributed to the fact that Otto does not deliver any information on the offered products and barely recognizes any product name. The usefulness examination based on task two displays a higher dispersion in Otto’s data. Participants perceived both Otto and Jana as generally less useful in comparison to Carla, while simultaneously providing a rather unanimous assessment thereof. Both tasks reveal a discrepancy of the VAs’ ability to offer useful information. Carla’s scores are quite high, but not all participants agreed on this fact. The differences between both tasks may result from the diversity thereof. Some participants might not have felt familiar with the mission of the 1st task and might have been more confident in the 2nd task, where they could ask open questions to the VA. However, as the VA’s usefulness is the most relevant aspect to use these services, the services should be improved to satisfy user needs.

In summary, Carla scored the best in both the heuristic as well as the user evaluation. She demonstrated a more natural understanding of the human language and offered a better developed understanding of users’ information need overall. This was recorded in the qualitative user statements as well, such as: “I could hold a real conversation with Carla, and even after some seemingly conversation derailing remarks, she was able to pick up where we left off” (survey data, May 2016). Only the fact that she is designed to under-

stand inflection form of words interferes with her otherwise well working auto correction. An obvious weakness of Otto and Jana appears to be the response calculation and the recognition of negation in user input, which in turn was some of Carla's best features, as a user recalled: "The information was accurate and it was delivered quickly. The questions were understood by the assistant and I was positively surprised" (ibid.).

5 Conclusion and future work

We analyzed and compared three VAs concerning their usefulness and user perception. Therefore, we conducted a heuristic analysis as well as a user study. First results show that users generally perceived one VA more positively, which might be due to the VA's stronger language processing abilities. However, user opinions were quite heterogeneous. It should also be mentioned that with our tests the entirety of these VAs could not be tested due to the lack of insight into their natural language processing technologies, meaning that there is potential for future improved evaluation procedures. It will be valuable to compare these first outcomes with the recorded user voices and searches that will give detailed insights into user needs and their behavior during the interaction with a VA. In the future, we will analyze the users' perception of the VAs' aesthetic and its influence on enjoyment and trust.

Acknowledgements

We would like to thank Kauz Linguistic Technologies for their support with creating VA assessment categories and all participants of our user study.

References

- Ahamed, S. I., M. Sharmin, S. Ahmed, M. M. Haque, and A. J. Khan (2006): Design and implementation of a virtual assistant for healthcare professionals using pervasive computing technologies. In: *E & I Elektrotechnik und Informationstechnik*, 123 (4), 112–120. [doi:10.1007/s00502-006-0335](https://doi.org/10.1007/s00502-006-0335)

- Chao, C.-Y., T.-C. Chang, H.-C. Wu, Y.-S. Lin, and P.-C. Chen (2016): The inter-relationship between intelligent agents' characteristics and users' intention in a search engine by making beliefs and perceived risks mediators. In: *Computers in Human Behavior*, 64, 117–125. [doi:10.1016/j.chb.2016.06.031](https://doi.org/10.1016/j.chb.2016.06.031)
- Chang, J. W., M. C. Lee, and T. I. Wang (2016): Integrating a semantic-based retrieval agent into case-based reasoning systems: A case study of an online bookstore. In: *Computers in Industry*, 78, 29–42. [doi:10.1016/j.compind.2015.10.007](https://doi.org/10.1016/j.compind.2015.10.007)
- Etemad-Sajadi, R. (2014): The influence of a virtual agent on web-users' desire to visit the company. In: *International Journal of Quality & Reliability Management*, 31 (4), 419–434. [doi:10.1108/IJQRM-05-2013-0077](https://doi.org/10.1108/IJQRM-05-2013-0077)
- Fong, T., I. Nourbakhsh, and K. Dautenhahn (2003): A survey of socially interactive robots. In: *Robotics and Autonomous Systems*, 42 (3–4), 143–166.
- García-Serrano, A. M., P. Martínez, and J. Z. Hernández (2004): Using AI techniques to support advanced interaction capabilities in a virtual assistant for e-commerce. In: *Expert Systems with Applications*, 26 (3), 413–426. [doi:10.1016/j.eswa.2003.09.012](https://doi.org/10.1016/j.eswa.2003.09.012)
- Keeling, K., P. McGoldrick, and S. Beatty (2010): Avatars as salespeople: Communication style, trust, and intentions. In: *Journal of Business Research*, 63 (8), 793–800. [doi:10.1016/j.jbusres.2008.12.015](https://doi.org/10.1016/j.jbusres.2008.12.015)
- Kuligowska, K., and M. Lasek (2005): Intelligent agents in the human-computer interaction. In: B. Kubiak and A. Korowicki (Eds.): *Information management. Proceedings of the 7th International Conference on Information Management – Human Computer Interaction*. Gdańsk. <https://www.chatbots.org/paper/intelligent-agents-in-the-human-computer-interaction/>
- Li, J. (2015): The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. In: *International Journal of Human-Computer Studies*, 77, 23–37. [doi:10.1016/j.ijhcs.2015.01.001](https://doi.org/10.1016/j.ijhcs.2015.01.001)
- McGoldrick, P. J., K. A. Keeling, and S. F. Beatty (2008): A typology of roles for avatars in online retailing. In: *Journal of Marketing Management*, 24 (3–4), 433–461. [doi:10.1362/026725708X306176](https://doi.org/10.1362/026725708X306176)
- Panetta, K. (2016): Gartner's Top 10 Strategic Technology Trends for 2017. <http://www.gartner.com/smarterwithgartner/gartners-top-10-technology-trends-2017/>
- Sarodnick, F., and H. Brau (2011): *Methoden der Usability Evaluation: Wissenschaftliche Grundlagen und praktische Anwendung* (2nd Ed.). Bern: Huber.
- Schumann, L., and W. G. Stock (2014): The information service evaluation (ISE) model. In: *Webology*, 11 (1).

Shiban, Y., I. Schelhorn, V. Jobst, A. Hörnlein, F. Puppe, P. Pauli, and A. Mühlberger (2015): The appearance effect: Influences of virtual agent features on performance and motivation. In: *Computers in Human Behavior*, 49, 5–11. [doi:10.1016/j.chb.2015.01.077](https://doi.org/10.1016/j.chb.2015.01.077)

Von der Pütten, A. M., N.C. Krämer, J. Gratch, and S.-H. Kang (2010): “It doesn’t matter what you are!” Explaining social effects of agents and avatars. In: *Computers in Human Behavior*, 26 (6), 1641–1650. [doi:10.1016/j.chb.2010.06.012](https://doi.org/10.1016/j.chb.2010.06.012)

Gamification Elements and Their Perception by Different Gamer Types

A Case Study for a Project Management Software

*Alexander von Janta Lipinski, Henrik Weber,
Ralph Kölle, Thomas Mandl*

University of Hildesheim, Germany
mandl@uni-hildesheim.de

Abstract

Gamification is being applied more and more to information systems. Many developers apply similar strategies for the gamification of tasks. A benchmark conducted in 22 mobile apps containing gamification elements showed reward points as the most frequent design element. In a pilot user experiment, we analyzed the user perception of three gamification elements including reward points in the context of project management. Reward points are seen as motivating; in contrast a leader board was viewed more critically. It was shown that the judgment of elements depends on the game personality type. The gamer type Killer had a medium and significant correlation with a positive evaluation of the gaming elements. Only the gamer type Achiever had a positive and significant correlation with a positive judgment of rewards points.

Keywords: gamification; human-computer interaction; project management; motivation

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 131–144.

1 Introduction

Games are playful activities, usually associated with fixed rules and done for enjoyment. Games can focus either on competitive or on cooperative behavior. Gamification is the introduction of game elements into other non-game related contexts. Gamification has been seen as a promising strategy in information systems to transfer positive aspects of games like enjoyment to other activities (Hamari & Koivisto, 2015). Gamification is mostly intended to motivate and involve users for work and other tasks like conserving energy. Positive effects can be higher enjoyment of users and better performance on the main task. Negative effects can include lower productivity through time spent on the game aspects and boredom if the game is not accepted or loses its attractiveness. Many studies on gamification in information systems have reported positive results (Thiebes et al., 2014; Hamari et al., 2014) but also doubts have been raised about the reliability of the research (Broer, 2014). People with different personality might enjoy different game designs or different elements of gamification. The acceptance of gamification may differ from person to person.

2 State of the art

Gamification was defined as “a process of enhancing a service with affordances for gameful experiences in order to support user’s overall value creation” (Huotari & Hamari, 2012: 19).

Improved motivation through gamification can be explained by the self-determination theory (Ryan & Deci, 2000). Humans are intrinsically motivated to fulfil their own basic needs. These include autonomy, competence and relatedness. All three of these factors are emphasized by games. Users act autonomously, gain game competence during the play and can connect to other players. Games and gamification elements need to target these basic needs. Often they also rely on extrinsic motivation through external rewards like points or physical benefits. The game related elements are not always easy to find for any task and consequently, a set of general gamification elements was developed. These have been classified by several publications. A coarse segmentation is provided by Nicholson (2013), who mentions reward-

based elements that are related to extrinsic motivation and game elements which are closely related to the primary task. An analysis by Thiebes et al. (2014) clustered the focus of the design strategies in system design, challenges, rewards, social influences and user specific. The three most often mentioned gamification elements which they name mechanics and dynamics were goals, achievements and a point system. In a literature overview, Hamari et al. (2014) identified the following so-called motivational affordances most often in research papers: points, leader boards, and achievements/badges. They also list rewards and challenge. It becomes obvious that not even the terminology in the area is yet unified and e.g. the relation between challenges, achievements and badges remain unclear. This is also the result of a literature analysis carried out by Seaborn & Fels (2015).

Gamification is widely researched, however, in business information systems there are still few applications. The overview article of Hamari et al. (2014) lists only two studies out of 24 which were analyzed. The application areas of gamification are manifold and range from environmental protection (Goldstein et al., 2008) to fostering exercise for patients (Stuart, 2014). Of specific interest are business applications. An ERP system was evaluated by Herzig (2012) in order to identify factors for technology acceptance. Test users were required to solve some business tasks within 15 minutes. They received virtual cash, which turned out to be very motivating. The prototypical design integrated 3-dimensional virtual worlds with gamification for typical SAP tasks so the effects can be based on both design strategies. Nevertheless, such studies can help to find appropriate TAM variants for gamified systems. An interesting experiment by Zagel and Bodendorf (2014) within the area of logistics for a task involved inventories at supermarkets. They showed that the addition of gamification elements to the inventory system increased the time that users spent on the task. However, at the same time the data quality increased. For the inventory application, a higher data quality is desired so that the increased cost is acceptable.

An increased data quality and better participation was also observed by Cechanowicz et al. (2013) within the area of market research. Their game elements included brief quizzes. In a five month long study with students with project management software, Schubert et al. (2014) found out that the motivation increased overall. However, the effect may decrease over time and comparative elements are seen as potentially demotivating and problematic. The authors stress that the context of the application needs to be carefully considered.

Apart from context, the acceptance of gamification elements may highly depend on the personality of a user. The personality can be described in many ways. For computer games, an influential classification of gamer types has been suggested by Bartle (1996):

- *Achievers* intend to follow the rules and win or achieve goals. They observe their progress and absorb goals set by the game design. Achievers like to step up in levels or accumulate points.
- *Socializers* seek a social experience and want to meet others and interact with other players. They tend to use communication tools and like communities.
- *Explorers* like to interact with the world in the game. They enjoy to discover new areas and to gain knowledge about the environment of the game. Explorers are curious and like new challenges.
- *Killers* are socially motivated and like to win over and dominate other players. They prefer competition and are eager to discover and learn new strategies and tactics to succeed.

A discussion of other categories is presented in Ferro et al. (2013). However, the categories suggested by Bartle (1996) are still popular in research. They are also used in our study.

The relation between personality and gamification design has also been subject to previous research. Ferro et al. (2013) provided a thorough analysis of literature on gamer types and personality. They suggested a plausible list of assignments of game elements to gamer types. However, their assessment is purely theoretical. Some practical suggestions for the appropriate elements for gamer types are also given by Cunningham and Zichermann (2011). The research on gamification is faced with methodological challenges. The effect of gamification on users during serious work tasks is difficult to measure in experiments. This effect cannot be researched by asking users out of context. Few long term studies, like the one over five months by Schubert et al. (2014) and one over two years by Hamari (2015), have been conducted.

3 Research questions and study design

Previous research has explored many facets of gamification in information systems. Many issues require more research. An important question is

whether gamification elements may be useful for anyone or whether different personalities judge the concept differently. Our hypothesis is that based on personality people evaluate gamification differently. There are several elements or design strategies for gamification. They also might be regarded quite differently. The perception of various strategies may also depend on the personality, which is coherent with the prediction of Ferro et al. (2013). If this hypothesis is proven correct, the design of gamification systems will need to take the personality of the prospective users into account or alternatively design consciously for a diverse audience. Another consequence is that questions on system design in general are perceived quite differently by people. We intended to create a realistic user experiment with real software involving interaction in realistic test tasks. During these tasks users are exposed to the gamification elements and were asked questions about these elements later. The perception of the elements needs to be correlated to personalities according to the gamer types.

4 Study

The experiment was carried out as a user test with the online project management software RedCritic, Tracker which includes gamification elements. RedCritic Tracker allows the typical project resource allocation tasks like creating tasks within projects, assigning them to workers and supervising the performance. Our study was intended to evaluate popular gamification elements. For that we took an empirical approach. In a benchmark, we collected data on 22 apps which included gamification elements. However, only 27% are fully available in Germany. Companies seem to be reluctant to introduce gamification elements into the German mobile market. Among these apps, 82% offered reward points which could be exchanged for real vouchers or products. Half of the apps connect to social networks systems.

Based on these results, we selected an existing application which included the most frequent gamification element, a reward point system. The system is not connected to social networks but implements networking functions with a project group. The project management software RedCritic Tracker (<https://www.redcrittertracker.com/>) can be used online. The main gamification elements implemented in RedCritic Tracker which were in the focus

of the study are reward points, badges and a leader board. All credits can be observed in a dash board by each user. The same elements can also be found in the Kudos system integrated into IBM connections.

These three gamification elements (reward points, badges and leader board) are also the most popular ones in research according to the literature review conducted by Hamari et al. (2014). They also appear in the empirically determined list assembled by Thiebes et al. (2014). Nevertheless, the two last mentioned studies used only scientific literature to identify gamification elements and strategies. Our approach was based on real systems that were found in app stores.

Instead of just presenting the software to the test users, a task scenario was created for the experiment. Within the scenario a company created the Web site for a coffee shop. Within this project, the test users were required to execute four typical tasks. The work tasks were designed to be typical within web design projects. Due to the fact that students are familiar with information systems, are frequently involved in project courses, are required to collaborate in electronic environments and often work outside the university in IT companies, the simple tasks seem quite natural.

The first task required the test users to change the position of the logo within a wire frame for the Web site and the second required changing the font type within the HTML code. These software related tasks had to be solved within the prototyping software Balsamiq Mockups for Desktop (balsamiq.com). The third task consisted of finding an expert for Android within the project team and of sending a message to her or him. A fourth task was filling out a questionnaire for the study and was also awarded with reward points.

A manager approving the tasks was simulated by a second person in another room. Applying this procedure, the test users could see their progress and the effect of their actions within the gamification elements. They earned credit expressed in reward points and advanced in the leader board. The test users were advanced students from two universities, who had gained some experience in team work during their studies. They received no incentive and formed a convenience sample. The test was carried out in a lab at the university in order to create an identical environment for all test users and to create a situation with no interruptions. All tasks were shown to the user in written form on the PC so that they all received identical information. On average, the test took 40 minutes and was followed by a questionnaire and a 5 minute

interview. The exposure of the users to the system was longer than in other studies (Herzig et al., 2012) but we did not aim at long term acceptance.

The questionnaire started with a few questions about project management and focused on the perception and evaluation of the gamification elements. For all three elements, the same set of questions was asked which consisted of three parts. In the first part, the perception of the elements was in the focus and participants were asked how much attention they dedicated to the gamification elements. The second part focused on the motivation and included Likert scales with statements dedicated to the specific elements. In the third part we used semantic differentials for three pairs of adjective in order to find out about the subjective evaluation of the elements. Semantic differentials are an established instrument to identify the attitudes of people.

Next, previous experience with video games and gamer type were gathered. Then, the participants were asked for their overall judgement of the gamification and the software. The questionnaire closed with socio-demographic data. The interview was aimed at collecting qualitative information and judgements of the participants.

5 Results

The user experiment was carried out by 20 people (12 male). They mostly liked project work with the RedCritic Tracker software (90%). Of the group, 60% liked video games and 40% did not like video games. However, 60% never or rarely play video games and only 20% play video games often. We assessed the gaming personality by the Bartle test based on a questionnaire. The gamer types are distributed as shown in figure 1. As shown some of the 20 individuals were assigned to two gaming types. The type Achiever predominates, however, our sample also contains many Killers. Overall, the gamification elements were perceived as positive. The two statements whether these are useful for project management and whether the gamification elements do not pose a barrier for work both receive predominantly agreement (15% fully agree, 60% rather agree for both).

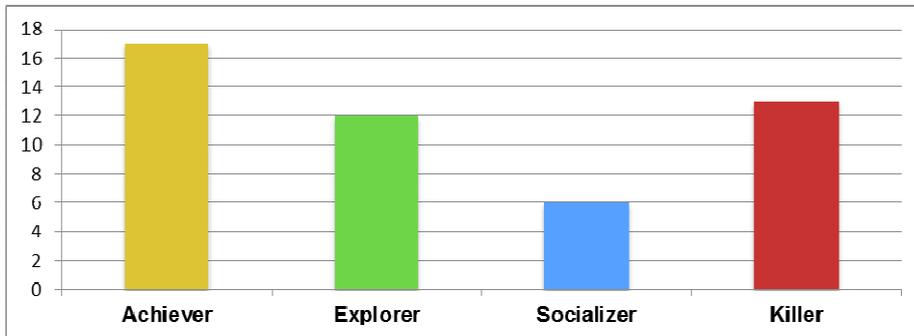


Fig. 1 Distribution of gamer types in test users

The study intends to get a differentiated judgment of the different gamification elements. Overall, the reward points were assessed most positively and the badges as most negative. The leader board received an assessment between the two other elements.

Among the three gamification elements the reward points were judged most often as positive. On a five point Likert scale for a statement about higher motivation, reward points as a gaming element received most very positive ratings and predominantly very positive and positive ratings (fig. 2).

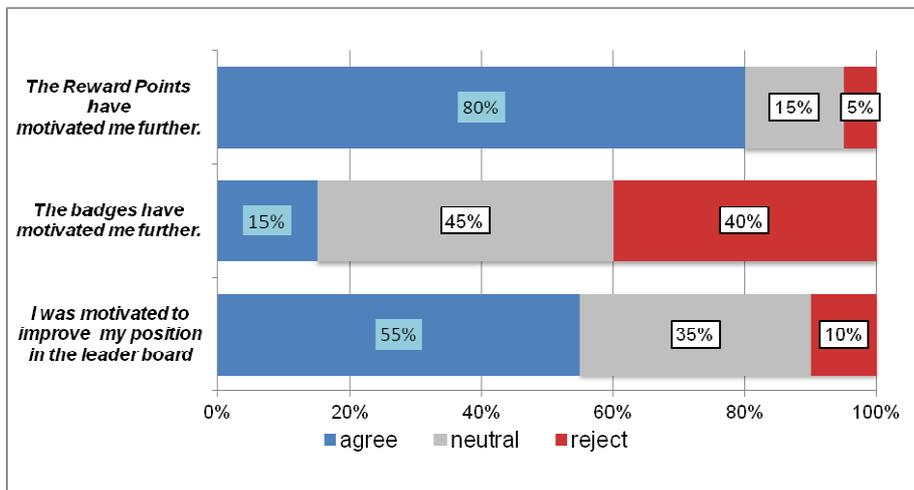


Fig. 2 Comparison of the three gamification elements

The perception of the badges was remarkably negative as figure 3 shows. Only concerning the curiosity of users, badges can get a neutral judgment on average. Subsequently, the users were asked how much attention they

had paid to the specific gamification elements. The answers are shown in figure 4. As shown, users judged that they have paid most attention to the reward points followed by the leader board. Least attention was given to the badges

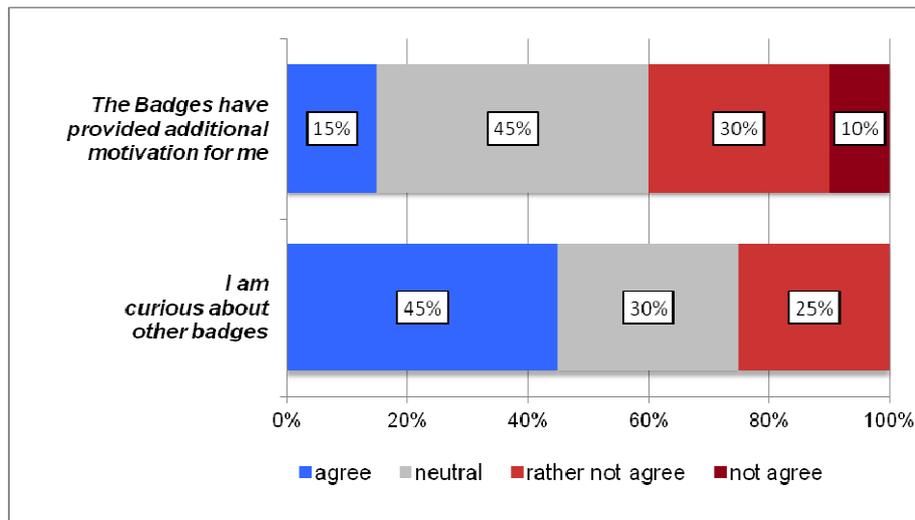


Fig. 3 Judgment of Badges

The results on perceived attention and the motivation provided by an element are in line with each other. Both questions lead to an identical ranking of the three elements under investigation.

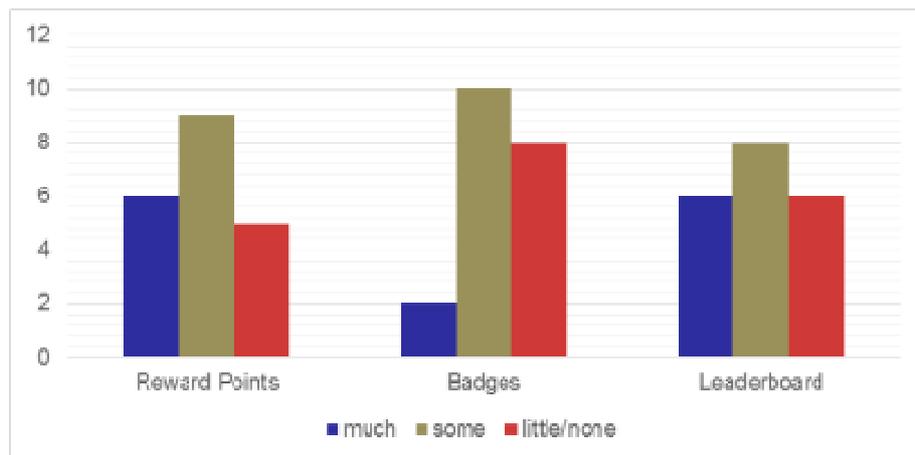


Fig. 4 Attention Dedicated to the Elements

In addition, the third way to compare the elements also leads to this ranking. In this part of the questionnaire, users were asked to rate the elements on a scale between bipolar adjectives. The semantic differential shows that the users associate the more positive adjectives with the reward points followed by the leader board as figure 5 shows. The differences between the three gamification elements are statistically significant for the first adjective pair (< 5%) and not for the second pair. For the third adjective pair, only the differences between the reward points and each of the other two elements are statistically significant (< 5%). The perception of the different gamification elements by the users is quite individual. We measured the gamer types of the 20 participants (cf. fig. 1) and searched for correlations to the judgments and preferences. Remarkably, the gamer type Killer correlated highly with an overall positive judgement of gamification elements in project management ($r = 0.54$, significant, error probability < 5%). The game type Achiever (and only this type) exhibits a medium correlation to the judgment of reward points as positive ($r = 0.55$, significant, error probability < 5%). Reward points might be mostly helpful for the personalities related to the game type Achiever.

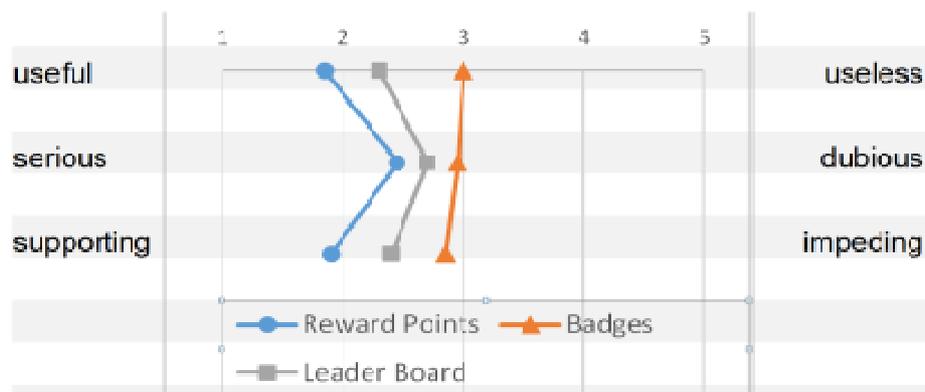


Fig. 5 Semantic Differential

In a theoretical mapping of game elements to player types, Achievers were assigned to badges and Killers to points (Ferro et al., 2013). These relations could not be found in our data. We further assessed the correlations between the answers of the users. The judgement of the leader board as supportive and the desire for competition also correlated positively ($r = 0.58$, significant, error probability < 1%). For all three elements, there is a medium

positive correlation between the judgment as useful and motivating. This is a plausible relationship.

After filling out the questionnaire, the participants had the opportunity to state their opinions in a brief interview. Most statements were again consistent with the results presented above. E.g. one person stated: “I liked the reward points a lot” Another participant said: “I like the idea that one gets rewarded for tasks, very good”. Despite the overall positive assessment, one participant showed a negative attitude towards reward points in general: “The quality of the tasks fulfilment could suffer if the focus lies on the points and not the task itself”. However, there is one exception. The leader board received many negative statements, e.g. “Respect among colleagues may be affected by the leader board” or “The leader board might be demotivating”. These statements show that designers need to be especially careful with the attitude of their specific clients and the business culture when introducing comparative elements like a leader board (for further details see Janta Lipinski & Weber, 2015).

6 Conclusions

We reported a user test as a pilot study with the gamified project management application RedCritic Tracker (RCT). Twenty students with previous experience in project management took part in this study. A detailed questionnaire was used to determine acceptance of game mechanics. Factors indicating acceptance, such as perceived increase of motivation as well as subjective impressions towards these game mechanics were examined.

The results of the user test suggest that the game elements examined in this study were mostly accepted by the participants in the test situation. Gamification elements were also viewed as motivational. Reward Points were accepted by most participants, Badges were least accepted, and the leader board was classified in the mid-position, even though it exhibits competitive characteristics which were often regarded as negative in qualitative statements. This result is consistent with the findings of Schubert et al. (2014) who conducted a five month study with German students and found opposition against public and comparative elements like the leader board. The perception of badges as most negative is surprising since they are seen as a very

positive element in the work of Hamari (2015). On the other hand, Schubert et al. (2014) asked after the exposure of students for five months which element people would remove first. The badges were named most often despite that fact that a leader board was also present and was highly criticized. More research on the type and design of badges and the context factors for their use seem to be necessary.

A small minority of the participants in the study uttered a negative attitude toward the game elements. Furthermore, our study shows also that the acceptance of game elements depends on other characteristics such as the urge for achievement and competition. The positive judgment of the gamification elements is in one case correlated to the gamer type of the participant (Achiever to reward points). This shows that not each user may accept gamification elements at the same level. The results also contradict theoretical assignments of game elements to personality types by Ferro et al. (2013). The authors associated the Achiever type with badges whereas we found a relation between Achiever and reward points. These results show that further empirical research is necessary both for the gamification of project management software and for other business information systems.

References

- Bartle, R. (1996): Clubs, Hearts, Clubs, Diamonds, Spades: Players Who suit MUDs. In: *The Journal of Virtual Environments* 1 (1). <http://mud.co.uk/richard/hclds.htm>
- Blohm, I., and J. Leimeister (2013): Gamification – Design of IT-Based Enhancing Services for Motivational Support and Behavioral Change. In: *Business & Information Systems Engineering*, 5 (4), 275–278. <http://aisel.aisnet.org/bise/vol5/iss4/6>
- Broer, J. (2014): Gamification and the Trough of Disillusionment. In: *Mensch und Computer 2014, Workshopband*. München: Oldenbourg, pp. 389–395.
- Cechanowicz, J., C. Gutwin, B. Brownell, and L. Goodfellow (2013): Effects of gamification on participation and data quality in a real-world market research domain. In: *Proceedings of Gamification 2013*. New York: ACM Press, pp. 58–65.
- Cunningham, C., and G. Zichermann (2011): *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*. Sebastopol, CA: O’Reilly Media.

- Ferro, L., S. P. Walz, and S. Greuter (2013): Towards personalised, gamified systems: an investigation into game design, personality and player typologies. In: *9th Australasian Conf on Interactive Entertainment, IE'2013, Melbourne, Australia, Sept 30–Oct 01*. New York, NY: ACM, Article No. 7.
- Goldstein, N., R. Cialdini and V. Griskevicius (2008): A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. In: *J. Consumer Research*, 35 (3).
- Hamari, J (2015): Do Badges increase user activity? A field experiment on effects of gamification. In: *Computers in Human Behavior*. [doi:10.1016/j.chb.2015.03.036](https://doi.org/10.1016/j.chb.2015.03.036)
- Hamari, J., and J. Koivisto (2015): Why do people use gamification services? In: *International Journal of Information Management*, 35 (4), 419–431.
- Hamari, J., J. Koivisto, and H. Sarsa (2014): Does Gamification Work? – A Literature Review of Empirical Studies on Gamification. In: *Proceedings of the 47th Annual Hawaii International Conference on System Sciences, Hawaii, USA*, pp. 6–9.
- Herzig, P., S. Strahringer, M. Ameling (2012): Gamification of ERP Systems – Exploring Gamification Effects on User Acceptance Constructs. In: *Multikonferenz Wirtschaftsinformatik*. Braunschweig, pp. 793–804.
- Huotari, K., and J. Hamari (2012): Defining Gamification: A Service Marketing Perspective. In: *Proceedings of the 16th International Academic MindTrek Conference*. New York, NY: ACM, p. 17.
- Janta Lipinski, Alexander von, and Henrik Weber (2015): *Akzeptanz von Gamification-Strategien im Projektmanagement: eine empirische Untersuchung*. BA Thesis, Univ. Hildesheim.
- Nicholson, S. (2013): Two paths to motivation through game design elements: reward-based gamification and meaningful gamification. In: *Proceedings of the iConference 2013. Texas*, pp. 671–672.
- Ryan, R., and E. Deci (2000): Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. In: *Americ. Psychologist*, 55 (1), 68–78.
- Schubert, P., L. Paulsen, and J. Hager (2014): Auswirkungen von Gamification in Enterprise Collaboration Systems. In: *Tagungsband Mensch & Computer*, pp. 3–14.
- Seaborn, K., and I. Fels (2015): Gamification in theory and action: A survey. In: *International Journal of Human-Computer Studies* 74, 14–31.
- Stuart, A. (2014): Exercise as therapy in congenital heart disease – A gamification approach. In: *Progress in Pediatric Cardiology*, 38 (1–2), 37–44.

- Thiebes, S., S. Lins, and D. Basten (2014): Gamifying Information Systems – A Synthesis of Gamification Mechanics and Dynamics. In: *22nd Europ. Conf Information Systems (ECIS)*.
- Williams, S. A., Terras, M. M., & Warwick, C. (2013): What do people study when they study Twitter? Classifying Twitter related academic papers. In: *Journal of Documentation*, 69 (3), 384–410.
- Zagel, C., and F. Bodendorf (2014): Gamification: Auswirkungen auf Usability, Datenqualität und Motivation. In: *Tagungsband Mensch & Computer*. München, pp. 15–24.

Session 4:
Information System Evaluation

Comparing Heuristic Walkthrough and User Studies in Evaluating Digital Appliances

Eva-Maria Meier, Patricia Böhm, Christian Wolff

Media Informatics Group, University of Regensburg, Germany
maria.l.meier@stud.uni-regensburg.de, {patricia.boehm,christian.wolff}@ur.de

Abstract

In this paper we present an empirical study comparing user studies and expert evaluations based on a specific set of heuristics for evaluating information appliances with a heuristic walkthrough. The study looks at an e-book reader as well as a digital music player. In the user study, question-answer protocols are used as means of intervention during the experiments. To gain insight into performance of the evaluation methods the identified problem sets were analyzed. Results for the thoroughness, validity and effectiveness are presented and compared with prior studies.

Keywords: intergenerational; knowledge sharing; generation; organizations; information and communication technology; tacit knowledge

1 Introduction

In the context of the internet of things (IoT), interactive technology is set to become more diverse than ever: People will interact on different platforms, using different devices and non-standard interface design. The promise of the invisible computer (Norman, 1998), rather than indicating less interaction, actually points towards ubiquitous media interaction. In this context, adequate methods for evaluating interactive systems are needed. In this paper,

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 146–157.

we present a comparative study of different approaches to evaluation information appliances: Starting from earlier research on heuristics of information appliances (Böhm, Schneidermeier & Wolff, 2014), two devices are evaluated using a heuristic walkthrough as well as user studies. The effectiveness of both methods is compared using thoroughness and validity as major criteria. In this study, we want to examine

- how good heuristic walkthroughs can be adopted using our set of heuristics for information appliances,
- gain information on thoroughness and validity of findings for the heuristic walkthrough using user studies as the methodological point of reference (comparison of evaluation methods) and
- compare the performance of the two methods for different appliances being examined.

The rest of this paper is organized as follows: In chapter 2, we give a short overview of the state of the art. In chapter 3, heuristics for evaluating information appliances are introduced. In chapter 4 we present metrics for comparing the output of different usability evaluation methods (UEMs). Chapter 5 gives an overview of the design of our study. Results are presented and discussed in chapter 6, and chapter 7 draws conclusions and gives a short outlook.

2 State of the art

In the last decades, starting from early human factors research and later becoming a major field of inquiry within computer and information science, human-computer interaction has developed a broad variety of methods for evaluation usability, defined in ISO DIN EN 9241-11:1998 (1998) as

“[t]he extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.”

The heterogeneity of evaluation methods is illustrated in handbooks such as Martin & Hannington (2012). While many methods have been well-established for a long time and can be organized along major oppositions like expert/user study, formative/summative evaluation, quantitative vs. qualitative

results, less is known about the relative effectiveness and efficiency of the methods themselves (cf. Law & Hvannberg, 2004).

For gaining insight into the performance characteristics of a usability evaluation method (UEM), Hartson et al. (2001) suggest using user studies as the reference point for comparison: Using real users for an application for device can generate a set of actual interaction problems against which other evaluation methods – in our case: heuristic walkthroughs based on our set of information appliance evaluation heuristics – can be matched.

3 Heuristics for evaluating information appliances

Picking up well-known sets of heuristics as put forward by Molich & Nielsen (1990) or Shneiderman (“eight golden rules”, Shneiderman, 1987) as well as relevant standards (ISO 9241-110:1998(en), 1998) and literature on heuristic evaluation, we have developed a set of heuristics for the evaluation of information appliances (Böhm, Schneidermeier & Wolff, 2014). This set consists of eight heuristic principles at the top layer, with additional sub-heuristics defined for each main category. The eight heuristics are defined as follows (translated from German):

1. *Consistency*: The appliance is designed consistently and conforms to applicable standards.
2. *Feedback*: Each interaction step should have immediate, appropriate, and recognizable feedback.
3. *Easy handling*: Handling should be as efficient as possible but at the same time give the user a feeling of being in charge.
4. *Error avoidance*: The design should take precaution that interaction errors concerning not-supported interaction, inadvertent interaction, or mix-up of function in interaction do not occur.
5. *Suitability for the task*: The device should provide the functions expected and needed by the user; the user interface should be designed to fit the tasks.
6. *Help and documentation*: In case of interaction problems and for helping to learn new functions, the device should provide adequate information.

7. *Self-descriptiveness*: Interaction for basic functions should be understandable without instructions or handbook usage.
8. *Flexibility*: Users with different competencies can use the device under different circumstances in everyday situations.

4 Problem comparison

The comparison of different UEM should be based on *common problem descriptions* (Lavery et al., 1997: 257 f.), consisting of a description of *context, cause, outcomes, breakdown in User's interaction, outcomes of the breakdown, outcome, solution*. This template was used for describing all problems found in this study. Problem descriptions from different evaluation methods can be compared and mapped if they are *consistently described in a common format* as suggested by Lavery et al. (1997).

In our study, we want to know how many actual or existing user problems can be found with a heuristic method. Thus, we need an initial set of “existing problems”. This can be generated by different methods (cf. Hartson et al. 2001: 390), we decided to generate a reference problem set by performing a user study. Given this set, the following metrics can be calculated (Sears et al., 2001: 388; Sears, 1997):

$$\text{Thoroughness} = \frac{\text{number of real problems found}}{\text{number of real problems that exist}}$$

$$\text{Validity} = \frac{\text{number of real problems found}}{\text{number of issues identified as problems}}$$

$$\text{Effectiveness} = \text{Thoroughness} \times \text{Validity}$$

For expert-based heuristic evaluations, the problem of false positives is well-known: Expert may “discover” usability problems that do not occur in users' actual interaction with a system or appliance. Therefore, *effectiveness* of a method means setting *thoroughness* in relation with *validity* as a rate for false positives (Sears, 1997: 213): The more problems described by experts turn out to be no “real” problems, the smaller the value for validity will be.

5 Study design

In Böhm, Schneidermeier and Wolff (2014) the following selection criteria for (information) appliances are defined: Small display, hard key as well as soft key interaction controls, and mobile context of usage.

5.1 Device selection

In our study, we have selected two devices which fulfil these elementary criteria (fig. 1):

- *Kobo Glo*
- *Apple iPod Nano*



Fig. 1 *Kobo Glo* e-book reader (left), *Apple iPod Nano* music player (right)

While *Kobo Glo* is a digital e-book reader, the *Apple iPod Nano* serves as a (MP3) music player. Selection criteria for the devices were as follows:

- In our previous study, we had evaluated a digital camera as well as a copying machine; in this study other/additional device types should be studied.
- At the same time, we intended to have devices with different assumed design/UX quality: Going along with the well-established UX design often found in Apple device we assume that the *iPod Nano* would have high usability ratings. At the same time, the opposite might be observed for the *Kobo Glo*, produced by a little known manufacturer with less experience in UX matters.

5.2 Task design

For developing adequate tasks for the user study that was performed to generate the reference set of “real” problems, a preliminary survey was conducted among possible users of both device types, e-book readers (20 questionnaires) as well as digital music players (27 questionnaires). For both device types, core tasks (e.g. reading; page navigation; searching / listening to music; shuffle function) as well as supporting tasks (setting markers; adjusting type size / video watching, surfing the internet) were identified. From this collection the actual tasks for the evaluation were generated.

6 Evaluation

The evaluation comprised two parts, the user study as well as the heuristic evaluation by experts. Both are briefly described below; we will not go into the details of particular interactions problems found in the study (cf. Meier, 2015, for an in-depth discussion of identified problems) as we want to focus on the aspect of method comparison here.

6.1 User study

In the user study, 20 test persons were presented with the tasks for the two devices. Following Nielsen, 2000, we assume that with 20 test persons, the actual amount of “real” problems can be approximated quite well. In the user study, test persons were recorded using a webcam. A moderator was present for conducting the experiment. In addition to the *thinking aloud* method, a *question-answer-protocol* was used for documenting problem situations (Grossman et al., 2009). Pre- as well as post-task questionnaires were used to document demographics and users’ experience with the device types as well as post-test ratings of the devices and their respective functionality and usability. All usability problems were documented using the template suggested by Lavery et al. (1997) along with a severity rating of the usability problem.

6.2 Heuristic evaluation by experts

Two experts were selected for the heuristics walkthrough based on the heuristics as introduced in chapter 3 above. Both experts have a background in usability engineering (information science / media informatics), with one expert having a junior level of experience while the other was already at a senior level (5+ years of UX experience). Using the test scenarios and heuristics as described above, both experts performed heuristic walkthroughs and documented their results. An explicit guideline for the heuristic evaluation was used in order to make sure that both experts followed a similar process in the heuristic walkthrough. A camera was used for documenting the usability problems found. For the problem documentation, the same template as in the user study was used.

7 Results and interpretation

The participating test persons in the user study consisted of 2 pupils, 7 undergraduate as well as graduate students, and 11 adult employed persons (age range 17–49 years). 20% of them possess an e-book reader, 65% a music player (digital media device). The following figures 2 and 3 show the amount of problems identified per user and device:

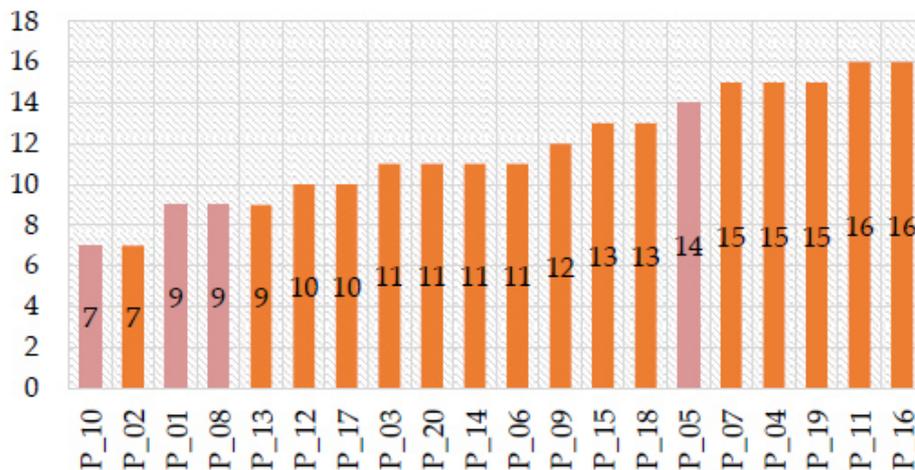


Fig. 2 Number of identified problems per User – Kobo Glo e-book reader



Fig. 3 Number of identified problems per User – Apple iPod Nano music player

In the post-task questionnaire, only few users showed willingness to buy the e-book reader (15%), while 70% would be willing to do so for the music player. The e-book reader was rated with 37.25 on average on the system usability scale (SUS) (Brooke, 1996, 2013), a fairly low value, while the music player reached a quite impressive value of 84.75 on average. The initial assumption of a broad difference in design and UX quality could be confirmed.

In the heuristic walkthrough, the experts identified the following number or problems (cf. tables 1 and 2, items with no problems found left out):

Table 1:

Number of identified problems per Expert – Kobo Glo e-book reader

Criterion	# problems for expert 1	# problems for expert 2
Consistency	4	13
Feedback	1	2
Easy handling	7	3
Error avoidance	8	13
Suitability for the task	3	2
Sum	23	38

Table 2:
Number of identified problems per Expert – Apple iPod Nano music player

Criterion	# problems for expert 1	# problems for expert 2
<i>Consistency</i>	0	10
<i>Feedback</i>	0	0
<i>Easy handling</i>	1	0
<i>Error avoidance</i>	3	2
<i>Suitability for the task</i>	1	0
<i>Help and documentation</i>	1	2
<i>Self descriptiveness</i>	1	0
Sum	7	14

In a next step, the problems identified by both experts were mapped onto each other using the common problem description format (Meier, 2015: 107 ff.). After this identification of overlaps, the overall expert problem set was constructed. Next, the problems sets for both UEMs were compared:

- *Kobo Glo*: 30 “real” user problems out of 38 overall problems could be mapped to 30 expert problems (out of 43 problems).
- *Apple iPod Nano*: 8 “real” user problems out of 13 overall problems could be mapped to 8 expert problems (out of 17 problems). For details of the problems and their mapping as well as a discussion of all problems identified by one type of UEM only, see Meier (2015: 113 ff.).

Finally, we have calculated the quantitative metrics as introduced in chapter 4 above:

$$\text{Thoroughness}_{\text{Kobo Glo}} = 30 \div 38 = 0.79$$

$$\text{Thoroughness}_{\text{Apple iPod Nano}} = 8 \div 13 = 0.62$$

$$\text{Validity}_{\text{Kobo Glo}} = 30 \div 43 = 0.70$$

$$\text{Validity}_{\text{Apple iPod Nano}} = 8 \div 17 = 0.47$$

Finally, effectiveness was calculated as the product of thoroughness and validity:

$$\begin{aligned} \text{Effectiveness}_{\text{Kobo Glo}} &= \text{Thoroughness}_{\text{Kobo Glo}} \times \text{Validity}_{\text{Kobo Glo}} = \\ 0.79 \times 0.70 &= 0.55 \end{aligned}$$

$$\begin{aligned} \text{Effectiveness}_{\text{Apple iPod Nano}} &= \text{Thoroughness}_{\text{Apple iPod Nano}} \times \text{Validity}_{\text{Apple iPod Nano}} = \\ 0.62 \times 0.47 &= 0.29 \end{aligned}$$

8 Discussion and outlook

Table 3 shows a comparison of these results with results from a previous study (Böhm, Schneidermeier & Wolff, 2014):

Table 3: Comparison of results with a prior study

	Camera	Copying Machine	<i>Kobo Glo</i> e-book reader	<i>Apple iPod Nano</i>
Thoroughness	0.77	0.79	0.79	0.62
Validity	0.91	0.79	0.70	0.47
Effectiveness	0.70	0.62	0.55	0.29

It becomes clear that results for the music player are much worse than for all other three devices. One might assume that the higher design quality – or the assumed higher design quality *as perceived by the experts* – plays a role in this outcome. The differences are more or less the same for both experts in the study.

Regarding the method, the combination of information appliance heuristics and heuristic walkthrough using a guideline worked quite well as the heuristics are more precise (only the top level is presented in chapter 3 above) than the more general heuristics discussed in the literature.

For the user study, using a question-answer-protocol as a means of documenting interventions by the moderator was successful in the sense that more tasks could be completed by the users. Finally, using precise templates for problem description had a steep learning curve in the beginning, but proved to be very helpful for problem identification and mapping.

The effects of (perceived) good interaction design quality for expert-based UEMs should be studied in more detail in the future.

References

- Böhm, P., T. Schneidermeier, and C. Wolff (2014): Heuristiken für Information Appliances. In: A. Butz, M. Koch, & J. Schlichter (Eds.): *Mensch & Computer 2014 – Tagungsband*. Berlin: De Gruyter Oldenbourg (pp. 275–284).

- Brooke, J. (1996): SUS: A “quick and dirty” usability scale. In: P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland (Eds.): *Usability Evaluation in Industry*. London: Taylor and Francis (pp. 4–7).
- Brooke, J. (2013): SUS: a retrospective. In: *J. Usability Studies*, 8 (2), 29–40.
- Grossman, T., G. Fitzmaurice, and Attar, R. (2009): A survey of software learnability: metrics, methodologies and guidelines. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM (pp. 649–658).
- Hartson, H. R., T. S. Andre, and Williges, R. C. (2001): Criteria for evaluating usability evaluation methods. In: *International journal of human-computer interaction*, 13 (4), 373–410.
- ISO 9241-11:1998(en) (1998): Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability. <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-1:v1:en>
- Law, E. L. C., and E. T. Hvannberg (2004): Analysis of strategies for improving and estimating the effectiveness of heuristic evaluation. In: *Proceedings of the third Nordic conference on Human-computer interaction*. ACM (pp. 241–250).
- Lavery, D., G. Cockton, and M. P. Atkinson (1997): Comparison of evaluation methods using structured usability problem reports. In: *Behaviour & Information Technology*, 16 (4–5), 246–266.
- Lazar, J., J. H. Feng and H. Hochheiser (2010): *Research methods in human-computer interaction*. Chichester: Wiley.
- Martin, B., and B. M. Hanington (2012): *Universal methods of design: 100 ways to research complex problems, develop innovative ideas, and design effective solutions*. Beverly, MA: Rockport Publishers.
- Meier, Eva-Maria (2015): *Heuristische Evaluation versus Nutzerstudie – Vergleich der beiden Methoden für die Evaluation von Geräten* [Heuristic Evaluation vs User Studies – A Comparison of Both Methods for Evaluating Information Appliances]. B. A. Thesis, Media Informatics Group, University of Regensburg, September 2015.
- Nielsen, J. (2000): Why You Only Need to Test with 5 Users. <http://www.nn-group.com/articles/why-you-only-need-to-test-with-5-users/>.
- Nielsen, J., and R. Molich (1990): Heuristic evaluation of user interfaces. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM (pp. 249–256).
- Norman, D. A. (1998): *The invisible computer: Why good products can fail, the personal computer is so complex, and information appliances are the solution*. Cambridge, Mass.: MIT Press.

- Sears, A. (1997): Heuristic walkthroughs: Finding the problems without the noise.
In: *International Journal of Human-Computer Interaction*, 9 (3), 213–234.
- Shneiderman, B. (1987): *Designing the user interface: Strategies for effective human-computer interactions*. Reading, Mass.: Addison-Wesley.

Development of a Benchmark System for Social Enterprise Software

Benchmark System and Visualisation for Analysing Personal Knowledge Behaviour

Wolfgang Semar, Elena Mastrandrea, Fabian Odoni

Swiss Institute for Information Research (HTW Chur), Switzerland
{wolfgang.semar, elena.mastrandrea, fabian.odoni}@htwchur.ch

Abstract

Network knowledge management in companies does not work without proactive motivation of their users. This paper describes the development of different benchmarks to assess users' performance and shows a novel approach to stimulate their willingness to actively share their knowledge in their collaborative work by the use of visualisations.

Keywords: network knowledge management; incentive system; motivation; measurement; benchmark system; data visualisation

1 Introduction

Well-established and popular social enterprise software like Atlassian Confluence, Microsoft Yammer and many open source projects show the high standard of computer-based support of knowledge communities and work groups. Growing competition makes knowledge an increasingly important success factor for enterprises. The resource-based view of creation, organisation, and use of intellectual capital is rated as an essential competitive issue.

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 158–163.

Since knowledge is often exclusively attached to interpersonal exchange, the cooperative and communicative aspect becomes highly important when it comes to knowledge management. In accordance with the paradigm of cooperative and communicative knowledge management, it is necessary to get over the dominating approach of knowledge warehouses and recognize the value added functions of electronic communication and interaction platforms for knowledge generation. The assumed supremacy of collaborative knowledge management is based on the productive exchange and sharing of knowledge among virtually connected groups, which balances knowledge asymmetry. It has been shown that the participants in such electronic systems need to be proactively motivated and supported (Schanz, 1999). We want to introduce such a system.

2 Developing a benchmark system for social enterprise software

A key method to support and encourage employee motivation is to allow them to understand how they and their colleagues deal with knowledge. To create this understanding, it is necessary to define appropriate benchmarks. The significance of individual benchmarks is limited without knowing the context, so there is the risk of an inadequate interpretation of the individual benchmarks. To avoid this, it is necessary to present the benchmarks by means of suitable visualisations. The lack of meaningfulness of individual benchmarks is countered by the combination of a selected set of benchmarks. It makes sense to link several factually related benchmarks to a benchmark system that describes the relationships and mutual effects of the individual benchmark. To develop benchmarks for knowledge management, the well-known basic systems of bibliometry, as presented by Havemann (2009), can be used, adapted and extended. To serve as an intrinsic incentive system such a benchmarking system must also provide the participants with appropriate feedback.

Starting at the basic functions in knowledge management tools, the following three basic variables for the development of benchmarks can be presented:

1. “Entity”: An entity is a user, or any form of an entry, a “like”, or other elements.
2. “Activity”: An “entity” is created by a certain activity by another entity. This can be coded as subject-predicate-object triple.

Examples:

User (entity) → opened (activity) → a blog (entity);

User → comments → blog entry

Blog entry → was commented → by a user

Space → has been opened → by a user

3. “Time”: Each “activity” occurs at a time

With the aid of these triples (entity → activity → entity) it is now possible to display various benchmarks. The activities can be differentiated by number (quantity) and by content (quality) and allows the development of different benchmarks.

To show the individual performance of users, automatically generated benchmarks are used. In order to set up the incentive/motivational benchmarking component, it is not helpful to use a hierarchic method, since not all benchmarks are mathematically related. The more useful approach is to have the measures in an order defined by subject and content criteria (Hummel, 2003: 555). Grob (2004: 50) suggests a benchmark system for LMSs (learning management systems) from which we borrow the benchmarks: coverage, relation, and time range that are registered on different levels: system level, group level, and individual level. Coverage is generated from measures like number of participants and entries and is given as absolute numbers (and sums). The combination of absolute numbers generates relation figures. They are shown as percentage or index numbers (Schwickert & Wendt, 2000: 8). Time range figures are derived from monitoring long-time user performance. By analysing timelines, changes in benchmarks can be identified. To create qualitative benchmarks, text analytical methods must be used. The quality of a contribution can also be inferred by the sensible combination of quantitative parameters. For example, the number of likes or comments to a post may be indicators of the quality of contributions. Additional metadata and tags set by the author or other users can be used to determine which author writes on which topic. By combining certain benchmarks, then it is possible to identify if a user is an expert on a specific topic.

3 Visualisation of benchmarks for social enterprise software

A visual aid in the form of a graphical representation can considerably simplify the interpretation of the benchmarks by the user. For the visual representation of communication relations, so-called knowledge maps are particularly suitable. Knowledge maps are virtual and represent immaterial data objects, which are not related spatially to one another. The knowledge maps are divided into the two categories “concept maps” and “associative maps”. The concept maps represent the subject areas in a specific arrangement and size. Concept maps are used in collaborative knowledge management, particularly to present the contributions of an individual actor, as well as to present the discussion context, and thus the knowledge distribution. The fact that concept maps are not particularly suitable for the visualisation of a hypertext-like system such as collaborative knowledge management lies in the fact that the edges are not explicitly represented. This problem is solved by associative knowledge maps (Däßler, 2002: 13–18). The associative maps visualise only objects and their object relationships. In associative maps, two types of associative structures can be distinguished, the “tree structures” and the “network-like structures”. The former only allow relationships between certain objects, while network structures in principle allow each object to be related to any other object. With this type of visualisation, it would be possible to represent the discussion structure of whole groups. The analysis and visualisation of the relationships in social networks is its own research field. The complexity of such networks is determined by the analysed characteristics, such as centrality in networks, grouping, distribution of roles, different relationships of the same set of actors, or the comparison of different networks. The fundamental measures to characterize the centrality of an actor in a community and the measures to characterize the importance of an actor provide an excellent basis for analysing the activities in a collaborative knowledge management tool (Dehmer, Emmert-Streib & Pickl, 2015; Cross & Parker, 2004).

The Atlassian Confluence tool is used as basic software. At present, 150 employees of a research and development department, which are distributed over six locations worldwide, share their knowledge through this tool. The goal is to analyse the communication and the written texts (in the form of wikis) in order to determine which person has what knowledge and how this

knowledge is exchanged between the employees. In doing so, the employees should be given a transparent feedback on their actions through visualisations and thus be motivated to pick up the knowledge of colleagues, but also to hand over their knowledge voluntarily.

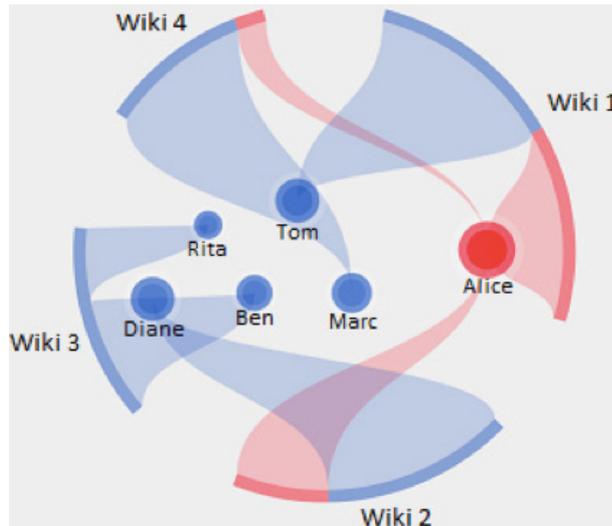


Fig. 1 Example of a wiki collaboration map for network knowledge management

Figure 1 shows the individual and the overall contributions for selected wikis. A node represents an author and an arc along the circumference the overall contributions to a specific wiki. Each section of the arc and corresponding chord that connects to an author's node displays their contributions to a given wiki. The larger the node, the more contributions that author has contributed. The red node represents the author who has contributed the most.

4 Future development

The comparison of the individual scores and making it visible to every member is also a strongly motivational momentum. It is also a proof of discourse control. It has to be kept in mind however that these benchmarks work on a quantity basis and do not reflect quality issues. To rate the quality of dis-

course objects, it is necessary to analyse content (intellectually and/or automatically). The first evaluation of the benchmark system showed that the benchmarks have to be refined and that advanced visualisations will be helpful. For the future, it is planned implement automatic text analysis in order to be able to better identify experts and thematic clusters. Furthermore, it should then be possible to analyse questions from users and forward them automatically to the relevant experts. In addition, the visualisation system will be extended by extensive interactive presentations. With the aim of motivating the participants to exchange their knowledge with their colleagues more and more, methods of gamification will be used.

References

- Cross, Rob, and Andrew Parker (2004): *The hidden power of social networks. Understanding how work really gets done in organizations* (2nd printing ed.). Boston, Mass.: Harvard Business School Press.
- Däßler, Rolf (2016): Visuelle Kommunikation mit Karten. <http://docplayer.org/4125712-Rolf-daessler-visuelle-kommunikation-mit-karten.html>
- Dehmer, Matthias, Frank Emmert-Streib, and Stefan Pickl (2015): *Computational Network Theory: Theoretical Foundations and Applications*. Wiley-Blackwell.
- Grob, Heinz Lothar, Frank Bensberg, Lofi Dewanto, and Ingo Düppe (2002): Controlling von Learning Management-Systemen – ein kennzahlorientierter Ansatz. In: Carstensen, Doris, and Beate Barrios (Eds.): *Campus 2004. Kommen die digitalen Medien an der Hochschule in die Jahre?* Münster: Waxmann (pp. 46–56).
- Havemann, Frank (2016): *Einführung in die Bibliometrie*. Berlin: Gesellschaft für Wissenschaftsforschung.
- Hummel, Thomas (2003): Quellen und Elemente von Informationssystemen des Controlling. In: Steinle, Claus, and Heike Bruch (Eds.): *Controlling – Kompendium für Ausbildung und Praxis*. Stuttgart: Schäffer-Poeschel (pp. 555–569).
- Schanz, Günther (1999): Motivationale Grundlagen der Gestaltung von Anreizsystemen. In: Schanz, Günther (Ed.): *Handbuch Anreizsysteme in Wirtschaft und Verwaltung*. Stuttgart: Poeschel (pp. 3–30).
- Schwicker, Axel C., and Peter Wendt (2000): Controlling-Kennzahlen für Web Sites. Arbeitspapiere WI, Nr. 8/2000, Lehrstuhl für Allg. BWL und Wirtschaftsinformatik, Johannes Gutenberg-Universität, Mainz.

Multilinguality of Metadata

Measuring the Multilingual Degree of Europeana's Metadata

Juliane Stiller¹, Péter Király²

¹ Berlin School of Library and Information Science
Dorotheenstr. 26, 10117 Berlin, Germany
juliane.stiller@ibi.hu-berlin.de

² Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen
(GWDG), Am Faßberg 11, 37077 Göttingen, Germany
peter.kiraly@gwdg.de

Abstract

Digital cultural heritage portals provide universal access to cultural objects and associated metadata originating in diverse countries and language milieus. Offering an equally heterogeneous audience access to this content is a challenging endeavour. To ensure accessibility for audiences with different linguistic backgrounds, it is crucial that the underlying metadata offers the same information in several languages. This paper presents the conceptualisation and implementation of a metric for measuring the multilinguality in the digital cultural heritage portal Europeana. For every field in each record across the entire collection, the level of multilinguality can be assessed. Quantifying the multilingual richness of data has significant benefits for increasing metadata quality, improving multilingual access to cultural collections and reaching multilingual audiences.

Keywords: multilinguality; measurement; big data; Europeana; digital libraries; metadata quality; evaluation

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 164–176.

1 Introduction

Metadata is at the heart of information systems intended to provide broad access to cultural heritage material which is often heterogeneous and published in diverse forms. One important information system offering access to digital cultural heritage is Europeana¹ – a platform that aggregates metadata from over 3,500 different memory institutions such as museums, libraries, archives and galleries. Originating in institutions from across Europe, the metadata is not only linguistically diverse but reflects the differing indexing practices of providing institutions. This heterogeneity is often a barrier to Europeana’s goal of offering broad access to its collection across languages for use and reuse.² The linguistic diversity of the describing metadata affects browsing, retrieval and display of the material and can be considered to be one dimension of metadata quality. High quality metadata ensures frictionless functionality; it is accordingly crucial to understand the factors that work to contribute to metadata quality. Research on this topic indicated that several different metrics (completeness, accuracy, timeliness to name a few) had previously been suggested; the topic of metadata multilinguality however, has not received much attention.

Multilinguality for Europeana means that a metadata record contains the same information in different languages and that values in certain fields are annotated with their respective language. It is evident, that records with more translations satisfy a greater number of functional requirements, e.g. search and access for a broad range of users, along with support for existing and anticipated functionality. The question is how this language diversity can be quantified to drive strategic decisions that improve multilingual functionality, such as display of content in users’ preferred languages, search across languages and semantic linking, in the long run.

This paper describes the conceptualisation and implementation of a quantitative measure of multilinguality for Europeana’s metadata. In the next section, section 2, related work is presented focusing on multilinguality in metadata as a dimension of metadata quality. Section 3 presents the concept of the multilingual score, while section 4 describes the implementation of the score

1 <http://www.europeana.eu/>

2 Europeana Strategy 2015–2020: http://pro.europeana.eu/files/Europeana_Professional/Publications/Europeana_Strategy_2020.pdf

into the Quality Assurance Framework (Király, 2015b). The paper ends with a discussion and prospectus for further developments.

2 Related work

2.1 Multilingual metadata in Europeana

Europeana aggregates over 54 million metadata objects³ from European cultural institutions. These metadata objects describe either digitized physical cultural heritage objects or born-digital material. Each object page in Europeana provides metadata describing that object, a thumbnail preview and a link to the owning institution's page for the object in question. Objects can be textual representations, images, audio or audiovisual content; of these, some 55% of the content consists of images.

For collections such as Europeana's, multilinguality is a key factor of fundamental importance. Not only is the described cultural content multilingual in itself, but so is its metadata. This linguistically diverse data is searched and accessed by an equally diverse audience from a wide range of countries and speaking many different languages. Offering information access independent of language is a challenging endeavour requiring solutions for the multilingual enrichments of metadata, features for multilingual user interactions such as search and browse functionality and an adapted graphical user interface design. Europeana has published a White Paper incorporating community input from various sources detailing all aspects which contribute to truly multilingual access provision (Stiller, 2016). Solutions Europeana has so far introduced to overcome language barriers are the automatic enrichments of metadata with multilingual vocabularies such as GeoNames (Manguinhas, 2016), language-independent access options such as colour search and the introduction of a knowledge graph for entities (Petras et al., 2017).

Automatic enrichment of metadata was evaluated (Olensky et al., 2012) with regard to its impact on retrieval. This process adds authority-type infor-

³ The size of Europeana is constantly increasing. On 24 January 2017, Europeana provided access to 54,217,972 million objects.

mation (person and place names, subject headings, date-time descriptors) from external vocabularies (such as VIAF, Wikipedia and GeoNames) that are often multilingual. On the one hand, the obtained results show that these processes need to be implemented with care to avoid negative impact on the user experience caused by incorrect enrichments, on the other hand, multilingual enrichments contribute to a higher visibility of documents in search results (Stiller et al., 2014), similar to the way query translation contributes to cross-lingual retrieval (Király, 2015a).

2.2 Multilinguality as a dimension of metadata quality

Attempts to improve multilingual information access cannot be undertaken without a holistic understanding of the multilinguality of metadata; however, there has heretofore been no methodology available for quantifying the multilinguality of a given record's metadata. Commonly accepted metadata quality dimensions and metrics do not include multilingual aspects – an astonishing omission given that access to data is one of the major motivators for improving and enhancing for data quality (Srivastava, 2011).

Eppler (2006: 71) identified up to 70 cited quality criteria for metadata; multilinguality, however, is not among them. Of course, multilinguality might be understood as a subcategory of other criteria such as completeness, accessibility and consistency. Such an approach, however, risks underestimating the multilingual problem, which remains a significant barrier for users who do not understand the language of the textual content – especially in digital collections (Chen, 2016: 17).

Taking an approach similar to Eppler's, Knight and Cowen (2005) reviewed 12 information quality frameworks in order to identify shared conceptualisations of information quality and derive a schema for assessing quality in the context of the World Wide Web. Again multilingual aspects are not mentioned. One can argue that the multilinguality of metadata is part of the user's perception of metadata belonging to the "subject criteria class" (Naumann & Rolker, 2000); here, a criterion such as "understandability" strongly depends on the language skills of a user. More objectively, the marking of language information in metadata can be determined and therefore measured. To the authors' knowledge, the only research which proposes a multilingual metric at all, measured the distribution of individual languages within a data collection based on language attribution (Vogias et al., 2013,

cited by Palavitsinis, 2014); the metric does not include whether a field's content is available in multiple languages.

To close this research gap, we have developed a model for measuring the multilinguality in metadata giving digital library administrators a means to assess their potential to reach multilingual audiences. This score and its implementation can further be used to support visualizations displaying patterns in the data which otherwise remain hidden. This development is part of the Quality Assurance Framework and Completeness measures of Europeana data (Király, 2015b). While the implementation described here focuses specifically on Europeana and the cultural heritage domain, the model itself is of course potentially applicable to other digital libraries.

3 A model for a multilingual score

To be able to calculate a multilinguality score for Europeana, one first needs to understand the organisation's information architecture and the potential multilingual dimensions which are reflected in it. The Europeana Data Model (EDM) is the metadata schema for the Europeana records (Isaac, 2013). It is based on RDF (Resource Description Framework), so a field value might have three types of values: literal (string or numeric), literal with language notation, and a resource identifier (URI) which points to another RDF statement. For example:

1. The value is a literal, e.g.
Subject: "Berlin Wall graffiti (writing)" .
2. The value has language annotation (the language should be encoded as an ISO-639 language code, here *en* for English, and *de* for German languages), e.g.
Subject: "Brandenburger Tor"@de , "Brandenburg Gate"@en .
3. The value is a resource identifier, pointing to a multilingual vocabulary such as GeoNames, e.g.
Subject: <<http://sws.geonames.org/2661886/>> .

These three potential formats of values populate the fields in an EDM record. They express a certain degree or level of multilinguality. According to the schema in table 1, we assigned scores to a field value ranging from 0 to 2.6 – the more multilingual information the higher the score.

For each field, the scoring in table 1 is used. If a field has a simple string value the scoring is 0, if the string value is marked with a language tag it gets a 1. If there are 2–3 different language tags the score 2 is applied, for 4–9 different language tags the score 2.3 and for more than 10 different language tags the score 2.6. Resource identifiers can be contributed by the Europeana data providers (those cultural heritage institutions, which share their records with Europeana), or they can be automatically added as the result of Europeana’s internal semantic enrichment process (Manguinhas, 2016). If a resource identifier is dereferenceable, the labels associated with the dereferenced entity are counted as though they belonged natively to the record; that is to say, the labels are considered to be in a sense ‘folded in’ to the record. On the other hand, if the identifier cannot be successfully dereferenced, then its contribution to the score is 0.

Table 1: Scores for field values with regard to multilinguality

Levels of multilinguality per field	Expressed in numbers
Missing field	NA
Text string without language tag (language not known)	0
Text string with language tag (language known)	1
Text string with 2–3 different language tags (language known with potential translations)	2
Text string with 4–9 different language tags (language known with potential translations)	2.3
Text string with more than 10 different language tags (language known with potential translations)	2.6

Obviously, in an ideal case the different language tags per field indicate translations of certain string values, but we are well aware that this is not always going to be the case. For some fields where one would expect a unique value (such as dc:title), we can assume that several labels with different tags indicate translations. For other fields where we often have several values (such as dc:subject), however, we cannot infer that the different instances are translations of each other. We are accordingly here simply counting distinct language tags, rather than translations per se.

Each field in a record is scored without a weighting. That means that a value in e.g. the dc:title field is not rated as more important than one in the dc:type field, and all fields are considered equal. This is the practice even in situations where one might expect some biasing – for example, with fields

such as `dc:title`, which will typically not contain links to a controlled vocabulary and thus tend to count lower than other fields.

3.1 Normalization

Normalization (scaling the scores to the range of 0 to 1) is considered beneficial for comparing, displaying and visualizing the data. One of the challenges is to determine how to normalize the score accurately. For now, a scaling of scores through

$$\text{normalizedScore} = 1 - 1/(\text{score} + 1)$$

is implemented. To avoid information loss during normalization, both scores – original and post-normalization – are stored and displayed.

3.2 Aggregating scores by instance, field, record or collection

The various approaches taken to score aggregation can be best illustrated by means of an example:

Subject field in record 1 with 2 instances of text strings with 2–3 different language tags each	Subject field in record 2 with 3 instances of text strings with 2–3 different language tags each
Instance 1: "table"@en, "tafel"@nl, "tisch"@de .	Instance 1: "flowers"@en, "bloemen"@nl, "blumen"@de .
Instance 2: "book"@en, "boek"@nl .	Instance 2: "cup"@en, "tasse"@de .
	Instance 3: "woman"@en, "frau"@de .
Sum: 4, Average: 2	Sum: 6, Average: 2

Here, we have several instances of the same field. Each one of the instances yields different scores. We calculate both the sum of the individual scores, and the average. Since there are more instances in record 2, it gets a higher sum, but the average will be the same.

For deeper investigation, the tool supports the retrieval of the aggregated scores and the list of values for the individual instances (cf. table 2).

The “instances” section contains the type and score of individual instances of fields; the “score” is for the final scores based on all instances. At the top and collection level, calculations are based on the scores of all in-

stances; only when inspecting individual records is information on particular fields preserved, displayed, and the score aggregation made explicit.

Table 2: Field level scores in REST API and in the web interface

REST API	Web interface
<pre>"Place/skos:altLabel": { "instances": [{"TRANSLATION": 2.0}, {"TRANSLATION": 2.0}, {"TRANSLATION": 2.0}, {"TRANSLATION": 2.0}, {"TRANSLATION": 2.0}, ... {"TRANSLATION": 2.40}, {"STRING": 0.0},], "score": { "sum": 20.40, "average": 1.85454545, "normalized": 0.649681 } }</pre>	<pre>instances ■ translation (2) ■ translation (2.40) ■ string (0) score: ■ sum: 20.40 ■ average: 1.85 ■ normalized average: 0.65</pre>

4 Implementation

The multilingual saturation score is implemented within the completeness measures of the open source Metadata Quality Assurance Framework: <http://144.76.218.178/europeana-qa/>. The framework is written in a modular way: the record level feature extraction and calculation of the score is written in Java using the Apache Spark framework, the statistical analyses were written in R and Scala⁴. The data was ingested from Europeana's OAI-PMH server⁵ and stored in Apache Hadoop's distributed file system as JSON files

⁴ Source codes are available from <http://pkiraly.github.io/about/#source-codes>, the workflow's details are described at <http://pkiraly.github.io/cheatsheet/>.

⁵ See <http://labs.europeana.eu/api/oai-pmh-introduction> for details of Europeana's OAI-PMH server. In order to make this research reproducible we published this snapshot

– one record per line. This way, the process could be easily parallelized and distributed over multiple processors and machines. The output consists of JSON files and PNG images. The web interface – written using PHP and d3.js – renders this output and provides rich navigation through, and interactive data visualizations of, the data.

Select dimension: grouped by

# records	Dataset	Minimum	Maximum	Range	Median	Mean	Standard deviation
238	35134 Parisienne de Photographie	35.4	333.4	298	113.05	116.1796	34.2511
647	5630 Deutsche Kinemathek	30.2	562.6	532.4	110.6	116.539	37.8709
335	20651 KU Leuven	52.2	321.4	269.2	102.4	110.0678	27.2582
969	1739 Bildarchiv Foto Marburg / Institut Mathildenhöhe Darmstadt	42.9	158.6	115.7	101	102.7493	17.7996
809	3190 Archiwum Muzyki Wiejskiej	67.2	124.5	57.3	98.5	92.2154	9.7659
411	14350 Universitätsbibliothek Leipzig / Digitaler Portraitindex	15	154	139	92.6	89.1179	20.9254
292	31 Germanisches Nationalmuseum Nürnberg / Digitaler Portraitindex	55	122.3	67.3	91.6	83.5129	18.5291
690	4695 MAK - Österreichisches Museum für angewandte Kunst / Gegenwartskunst	13	195.9	182.9	91.1	89.5043	18.2391
1731	184 Cinéma-thèque Royale de Belgique	27.8	309.1	281.3	84.4	91.4761	42.0275
1031	1388 www.esbifky.cz	40.5	129	88.5	82.5	80.9622	15.9246
380	3 LWL-Museum für Kunst und Kultur (Westfälisches Landesmuseum) / Digitaler Portraitindex	74.9	98.6	23.7	81.9	85.1333	12.1763

Fig. 1 Cumulative score of multilingual saturation per data providers, ordered by the median values

On the main page one can see the aggregated statistics for each dataset or data provider (fig. 1). By means of the drop-down menu, the multilingual saturation can be displayed for each field, plus the cumulated one showing the sum, the average or the normalized average of the multilingual saturation score. The table shows basic statistics, such as record count, minimum, maximum, mean, median values, range and standard deviation. The table is sortable, allowing ready exploration of the data; such an approach is particularly useful in identifying outliers that arise from data problems.⁶

The same information is available on a heatmap visualization (cf. fig. 2). Each dataset/provider in Europeana is represented by a square in the heatmap. When clicking on an individual square, i.e. data set/provider, statistics for this data set and its constituent fields are shown. All fields starting with “multilingual saturation” are of interest here.⁷

(created in the end of 2015, containing 46 million records, 1755 files, 420 GB in total) under this persistent identifier: <http://hdl.handle.net/21.11101/0000-0001-781F-7>.

⁶ An example of the sum of the multilingual saturation of the dc.title field can be found here: http://144.76.218.178/europeana-qa/?feature=saturation_sum_proxy_dc_title&type=data-providers.

⁷ One example is the dataset of the Rijksmuseum, the relevant score can be found from row 147 downwards. <http://144.76.218.178/europeana-qa/dataset.php?id=51&name=Rijksmuseum&type=d>

On the level of record investigation, you can see in the table “analysed metadata fields”, how the values were scored and what the score is for each field.

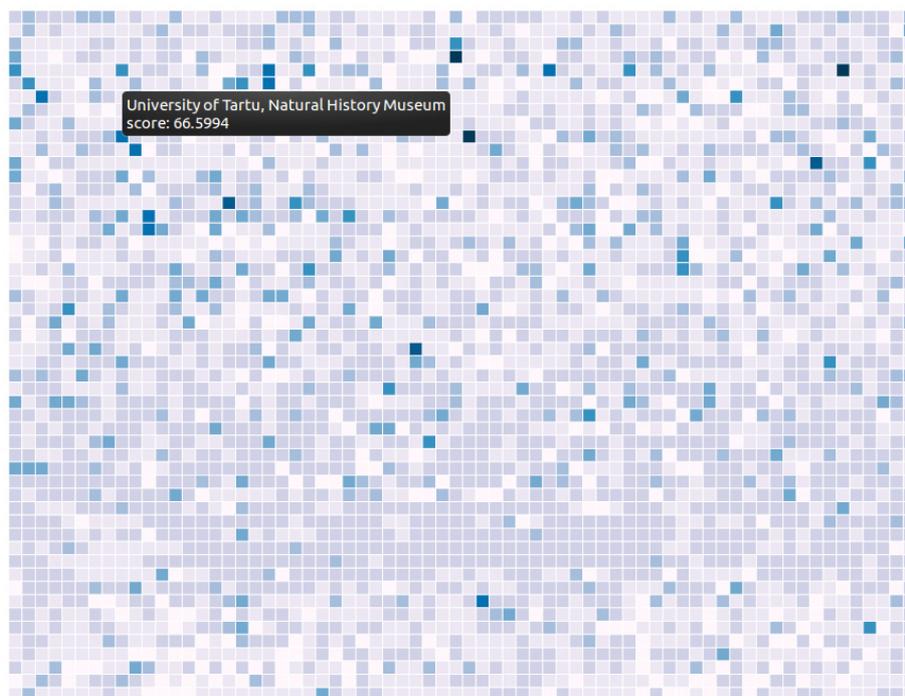


Fig. 2 Cumulative score of multilingual saturation per data provider as interactive heatmap: the darker the colour, the higher the score. The squares are linked to in-depth analyses of the collection.

This score is part of the completeness measures and should not be understood as an independent score. All fields are taken into account but the multilingual measure is bound to the completeness metric. So a missing field should not harm the multilingual score, but be reflected in the completeness score.

5 Discussion and further development

In this paper, we introduced a score for measuring the multilinguality of metadata. This metric assesses the values of different fields and describes the fields' multilingual variety or richness. This is a first attempt in quantifying multilingual information in metadata in digital collections. It is intended that this will in turn assist in improving the quality of metadata in this regard and raise awareness of the importance of multilinguality for information access. The different aggregation levels of the score across a field, such as dc:title, over the whole collection help to reveal data quality problems – the statistical analysis of the information can be leveraged with the several visualizations offered.

Improvements to the score will be made in future by reviewing and comparing collections, their resulting scores, the representation of the scores and the visualizations. Additionally, the occurrences of multilinguality in metadata will be linked to the user experience and to their impact on search, browse and other functionalities in the portal. By doing this, the score can be harnessed to its full potential.

Of special concern is the distinction between the multilingual potential of the metadata that is submitted by the providers and the multilinguality Europeana is able to add to the data automatically. Currently, the score reflects the multilinguality of a metadata record in a modified version of what the data providers submit – that is to say, after Europeana has optionally added information from external semantic vocabularies in the ingestion process. Since these are multilingual data sources, this process improves the overall score, obscuring the multilingual character of the original record. In future, we want also to measure the original records and determine the multilinguality of the record during different stages of the ingestion process. This will help identify strategies to exploit the multilingual potential of data more fully.

Acknowledgements

This research was conducted under the aegis of the Europeana Network's Data Quality Committee (DQC)⁸ formed early in 2016. We would like to thank all members of the DQC for their contributions to this research. We would also like to thank Timothy Hill from the Europeana Foundation for his feedback and the

⁸ <http://pro.europeana.eu/page/data-quality-committee>

review of this paper.

References

- Chen, J. (2016): *Multilingual Access and Services for Digital Collections*. Santa Barbara, CA: Libraries Unlimited.
- Eppler, M. J. (2006): *Managing information quality. Increasing the value of information in knowledge-intensive products and processes*. Berlin, New York: Springer.
- Király, P. (2015a): Query Translation in Europeana. In: *The Code4Lib Journal*, 27. <http://journal.code4lib.org/articles/10285>
- Király, P. (2015b): A Metadata Quality Assurance Framework. <https://pkiraly.github.io/metadata-quality-project-plan.pdf>
- Knight, S., and J. M. Burn (2005): Developing a framework for assessing information quality on the World Wide Web. In: *Informing Science: International Journal of an Emerging Transdiscipline*, 8 (5), 159–172.
- Manguinhas, H. (2016): Europeana Semantic Enrichment Framework. Documentation, Europeana. <http://bit.ly/2ijmKYM>
- Naumann, F., and C. Rolker (2000): Assessment Methods for Information Quality Criteria. In: *Proceedings of the 5th International Conference on Information Quality* (pp. 145–161).
- Olensky, M., J. Stiller, and E. Dröge (2012): Poisonous India or the Importance of a Semantic and Multilingual Enrichment Strategy. In: *6th Research Conference, MTSR 2012, Cádiz, Spain, November 28–30, 2012*. Proceedings, Berlin: Springer (pp. 252–263).
- Palavitsinis, N. (2014): *Metadata Quality Issues in Learning Repositories*. PhD thesis. Universidad de Alcalá. <http://dspace.uah.es/dspace/handle/10017/20664>
- Petras, V., T. Hill, J. Stiller, and M. Gäde (2017): Europeana – a Search Engine for Digitised Cultural Heritage Material. In: *Datenbank Spektrum*. [doi:10.1007/s13222-016-0238-1](https://doi.org/10.1007/s13222-016-0238-1)
- Srivastava, M. D. (2011): *Metadata creation in digital libraries*. Delhi: Pacific Publication.
- Stiller, J. (Ed.) (2016): White Paper on Best Practices for Multilingual Access to Digital Libraries. Europeana White Paper. http://pro.europeana.eu/files/Europeana_Professional/Publications/BestPracticesForMultilingualAccess_whitepaper.pdf

- Stiller, J., M. Olensky, and V. Petras (2014): A Framework for the Evaluation of Automatic Metadata Enrichments. In: S. Closs et al. (Eds.): *Metadata and Semantics Research: 8th Research Conference, MTSR 2014, Karlsruhe, Germany, November 27–29, 2014*. Proceedings (pp. 238–249). Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-13674-5_23
- Vassilakaki, E., and E. Garoufallou (2013): Multilingual Digital Libraries: A review of issues in system-centered and user-centered studies, information retrieval and user behavior. In: *International Information and Library Review*, 45 (1–2), 3–19. <https://doi.org/10.1016/j.iilr.2013.07.002>
- Vogias, K., I. Hatzakis, N. Manouselis, and P. Szegedi (2013): Extraction and Visualization of Metadata Analytics for Multimedia Learning Object Repositories: The case of TERENA TF-media network. In: *Workshop on Learning Object Analytics for Collections, Repositories and Federations, 9 April, 2013*. <https://www.terena.org/mail-archives/tf-media/pdf547CE3IKFt.pdf>

Text Mining for User Query Analysis

A 5-Step Method for Cultural Heritage Institutions

Anne Chardonnes, Simon Hengchen

Université libre de Bruxelles, Belgium
{anchardo, shengche}@ulb.ac.be

Abstract

The recent development of Web Analytics offers new perspectives to libraries, archives and museums to improve their knowledge of user needs and behaviours. In order to dive into the mind of their end users, institutions can explore queries from a digital catalogue. However, a manual exploration demands a major time commitment and only leads to limited results. This paper explores how text mining techniques can help automate the analysis of large volumes of log files. A 5-step methodology including clustering is illustrated by a case study from the State Archives of Belgium.

Keywords: user query; logs analysis; information retrieval; text mining; cultural heritage

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 177–189.

1 Introduction

In a context of budget cuts¹ and an increasing impulse to reach wider and more diverse audiences, cultural heritage institutions require a better understanding of user needs. Moreover, as expectations regarding content and services have grown with the evolution of the Web, it is essential for the institutions to understand the changing behaviour of users (Showers, 2015).

Beyond methods institutions developed to monitor the way collections are used *in situ*, from rare medieval maps to the copies of Anna Karenina, they have to imagine ways to monitor online activities. Web Analytics offer a new quantitative method to understand user behaviour. This unbiased observational data represents a unique opportunity to dive more precisely into the mind of end users through requests they made into the catalogue search engines (Grimes, Tang & Russel, 2007). From this approach, libraries, archives and museums can gain three kinds of insights.

First, by gathering statistical data, they can identify the types of content which turn out to be the most popular among their users. This information can be used as priority indicators regarding acquisition, cataloguing and digitisation of collections in order to meet user demands better. Secondly, user queries can be explored to better understand end users' behaviours and improve users' experience. Thirdly, they can discover the way users formulate what they seek and thus adapt the collection's metadata.

It should be noted that manual analysis involves a major time commitment for limited results. Indeed, it is difficult to obtain reliable numbers on the popularity of a topic, considering changes in the ways of writing queries. For example, it would be very time consuming to identify and reconcile manually every query related to *registres paroissiaux* (parish registers) with variant spellings (singular, plural, with space or separated by a "+", etc.). This is the kind of task that can be semi-automatically performed via text mining techniques, leading to potentially more precise information.

This paper explores the application of text mining techniques for user query analysis. With the help of a concrete case study from the State Archives of Belgium, we illustrate the possibilities but also the limits of using

¹ This was, for example, the case with the European Commission that not longer provides funding for the metadata creation since 2008 (van Hooland, Vandooren & Méndez Rodríguez, 2011).

text mining techniques in a cultural heritage setting. The article is constructed as follows. After this introduction, the second section presents an overview of existing works on user queries. The core of the paper consists of a 5-step method including clustering to analyse user queries, which is illustrated by our case study. The paper ends with results, discussion and future work.

2 Related works

Studying user queries and access paths to digital content in a cultural heritage context involves concentrating the attention on different research areas. At present, the way users search in digital catalogues is influenced by research conducted in other contexts. A well-known example is the Google search environment, which initiated the concept of the “Googlized” library patron (Woods, 2010).

Studies based on search engines are numerous and characterised by very different angles of approach: user information-seeking behavior, analysis of failed queries, domain knowledge or multilingual issues. Moreover, this field of research evolves rapidly, in parallel with the sophistication of search engines and the rise of new computational methods. For example, Grimes, Tang and Russell (2007) reviewed three sources of data that can be used to improve performance on the goals of a search engine: the field study, the usability study data and the raw query log. While noting the rare opportunity to gather information without disturbing the user, they underline the fact that the logs are not sufficient to measure the why, only the how and the what. Three years later, Kathuria, Jansen, Hafernik and Spink (2010) published a paper aiming to automatically classify the different user intents behind web queries.

In the cultural field, different methods have been used to investigate user search experience. To plan an academic library website redevelopment, Chase, Trapasso and Tolliver (2016) conducted a usability test. Though their study leads to actionable findings, such as a need for more support and information literacy instruction for students, it does not give much knowledge about the content of the queries themselves. More generally, Zavalina and Vassilieva (2014) observed that, among all the studies published by large scale digital libraries they reviewed, only a minority examines the content of

the user search queries. Ceccarelli, Gordea, Lucchese, Nardini and Tolomei (2011) analysed query logs to enhance the usability of the Europeana Portal and to develop assistance functionalities such as a query recommender system. Dijkshoorn et al. (2014) used log files from the Rijksmuseum to combine user queries with external vocabularies published as Linked Data, in the attempt to diversify search results. However, in both cases, no text mining methods appear to have been used on user queries.

By contrast, in the context of her analysis performed on the IMLS Digital Collection Registry transaction log dataset, Zavalina (2007) described how they first processed the queries (truncating plural, grouping together correct and misspelled versions of the same word, excluding stopwords such as prepositions, etc.) and then worked with a controlled vocabulary. However, the corpus contained less than 1 000 queries and apparently the whole process was done manually, including the extraction of all query strings. This last example illustrates the potential of text mining in this context: using a script to semi-automatically process the data could save time and be applied to large datasets. The workflow reported in this article seeks to bridge this gap by describing in detail each step and illustrating its utility.

3 Method

The present method aims at splitting the whole analysis process in five concrete steps. Since cultural heritage institutions have limited human resources at their disposal and are sometimes forced to outsource some technical tasks, special attention has been paid to provide open source solutions,² adaptable to various contexts. Regardless of the software architecture behind the search engine of a digital catalogue, our method allows every potential user to retrieve and exploit user queries in a semi-automated manner.

² The code is freely available online: <https://github.com/anchardo/PGCC>.

3.1 Collecting

The first step of the method is to actually collect the user-entered data. In the case of Web Analytics, three different possibilities exist to collect data: web log files; query parameters and JavaScript functions. Web log files contain raw data recorded automatically by web servers; query parameters composing URLs can be provided to analytics tools such as Google Analytics, which will retrieve and store them automatically; JavaScript functions can be used to capture and store the queried terms if they do not appear in the URL. Regardless of the method chosen, once the data has been collected and stored, it becomes processable.

3.2 Parsing

Within the context of Web Analytics, parsing is the task of reading through an URL and selecting, according to predetermined rules, the terms of the query and, when available, the advanced search parameters. This task, requiring possession of web log files, proves to be essential when the aim is to carry out an in-depth analysis.

Since all websites are different, as well as the structure of the corresponding URLs, the parsing method should be adapted for each unique website. Nonetheless, what can sound like a daunting task is made easy by the use of regular expressions (the so called “regex”)³, which can be customised to extract the information needed from any URL.

3.3 Grouping

Data analysis tools provide powerful functionalities to group and aggregate data, which will be useful in such a context. Thus, queried terms entered only once during a visit may still appear several times within the next URLs (for example when the visitor consults the results pages). Counting all these occurrences indifferently would skew study results such as the most searched terms. Grouping functionalities offers the possibility of applying an approximate but consistent method consisting of keeping each query at the

³ Regex is a text pattern following a specific syntax, which helps to find any string of characters matching that pattern within a text (Goyvaerts & Levithan, 2009).

most once per visit or per visitor, which helps to have a bird's eye view of the most popular search queries.

3.4 Cleaning

At this stage, user queries have been extracted and stored in tabular files. Before being analysed, they still need to be “cleaned”. The goal is to be able to associate similar string of characters despite superficial differences (Manning, Raghavan & Schütze, 2008), otherwise statistics could be skewed.⁴

That step can consist of various operations: to trim whitespaces – which means removing unnecessary spaces from the left-hand and right-hand sides of strings –, to convert all characters to lowercase or to replace special characters. Another process is tokenisation, which leads to chop each of the words composing a query, if there is a need to analyse them in an isolated manner. Finally, one could want to go further focussing on grammatical differences via stemming and lemmatisation, to reduce various forms of a word to a common base form (ibid.). Types of operations carried out during that step depend mostly on the data set and the needs of the analysis. The most important is to be consistent, i.e. to pre-process the whole data set the same way.

3.5 Clustering

This last step is designed to tackle different issues. First, user-entered data is always prone to errors. Secondly, the spelling of proper nouns is known to have evolved throughout the ages. Thirdly, the cleaning process only groups quasi perfect matches. It is therefore worthwhile to implement an additional function which seeks to cluster similar-yet-different terms. That task, called clustering, can be completed by various algorithms. The choice of the algorithm has to be made according to the data set particularities.

⁴ If user queries were not pre-processed, this could, for example, lead *Jan van Nijlen* and *Jan Van Nijlen* (capital “v”) to be considered as different queries.

4 Case study

The selection of our case study was guided by these three steps: find an institution willing to share its log files; obtain a consistent and large data set to test text mining techniques under realistic conditions; study user queries in a multilingual context.

Our active participation in a research project involving three national cultural heritage institutions of Belgium⁵ facilitated our access to raw data. Our final choice fell on the State Archives of Belgium since they also fulfil the two other conditions, with a four-language online catalogue and more than 175 000 visits per month.⁶ Moreover, until now, no research had been conducted by the institution on user queries, despite it being an important issue.

The online presence of the State Archives consists of an informational website and a digital catalogue⁷ called *Search*, whose user queries will be extracted as the raw data of our study. *Search* provides access to the database of the State Archives collection, which gathers archival heritage from the National Archives in Brussels and 18 repositories throughout the country. The scope of our study will focus on the main search engine, which is supplemented by two other devices enabling search by person or by producer.

4.1 Collecting

Among the three main possibilities to collect user queries (web log files, query parameters and JavaScript), we choose the first one, which allows us to gather additional information.

Search, the digital catalogue from the State Archives of Belgium, is tracked via a Piwik instance.⁸ A web server automatically records log files containing raw data such as the web pages users request or the IP addresses

5 The Maddlain project, a research project with the aim of modernising digital access to the collections of the Royal Library of Belgium, the States Archives of Belgium and the CegeSoma: <http://www.maddlain.iminds.be/en/>.

6 Numbers based on the average monthly visits between January and June 2016, which amounts to 176 510 visits, realised by an average of 44 795 unique visitors.

7 <http://search.arch.be/en>

8 The Piwik integration has been carried out by imec, a research institute linked to the University of Antwerp and the Gent University.

they are requested from. It has to be noted that those files are enormous⁹ and filled with data of no interest for such a study. In order to tackle performance issues due to that large amount of raw data and to avoid being forced to use more complex infrastructures, we used two strategies to reduce the amount of data to store locally. First of all, we reduced the survey period to six months (from 1 January to 1 July 2016). Second of all, we used a text filter¹⁰ to keep only URLs containing user queries.¹¹

4.2 Parsing

In order to handle our data set and apply advanced processing methods, we selected an open-source interactive programming environment: the Jupyter notebook.¹² Within that environment, we wrote a script in Python. Via three rather brief functions,¹³ the queries written by users can be extracted,¹⁴ as well as advanced parameters such as language preferences or filters based on time period or location of deposits.

4.3 Grouping

Within the Jupyter notebook, we used Pandas – an open source Python package – and more precisely its “group by” functionality to keep only one occurrence of each keyword(s) entered by a user during their visit.

In order to do so, we created a DataFrame containing three columns, based on informations parsed from the URLs: the ID of the visit, the queried

9 The intensive usage of the site (approximately 300,000 pageviews/day, mainly pages containing digitised genealogical sources) produces more than 2 GB of log files per month.

10 A regex inserted within the SQL LIKE operator.

11 Less than 1% of the URLs that correspond to the 300 000 pageviews/day contain search terms entered into the main search engine (the majority of users arrive at digital sources by means of links rather than by making use of search functionalities).

12 <http://jupyter.org/>

13 For the sake of brevity, the code is not presented here, but it is available online: <https://github.com/anchardo/PGCC>.

14 Here is an example of an URL containing the specific search terms (“belgium map”) entered in the search engine: [http://search.arch.be/en/zoeken-naar-archieven/zoekresultaat/index/index/zoekterm/belgium map/lang/en](http://search.arch.be/en/zoeken-naar-archieven/zoekresultaat/index/index/zoekterm/belgium%20map/lang/en).

term(s) and the time. Time has been used here as an HTTP request identifier: it eases the distinction of several occurrences of the same query (strings are identical) during the same visit (the ID numbers are identical). Using the “group by” operation, we have been able to group rows containing the same queried term(s) and, combined with the “count method”, we obtained in a single row the visit ID, the user query and the number of occurrences per visit.

4.4 Cleaning

User queries were stored in a CSV file. In order to ensure a better understanding of the queries, the decision was made to not parse queries into separate words. However, it appears that there are “hidden duplicates”. In order to clean our data and unmask these similar queries, we decided to convert all characters to lowercase, to replace each character which is not alphanumeric (for example “+”) by a space, to replace special characters (for example, replacing “é” by “e”) and to trim whitespaces. Thus, an original user query such as “matheus+De+Vuyst” becomes “matheus de vuyst”. In order to identify most requested content, the grouping operation was processed once again after this cleaning step.

4.5 Clustering

The algorithm which seems most appropriate for our context is based on the Jaro-Winkler distance (Winkler, 1990), a variant of the more common Jaro distance (Jaro, 1989). This method aims at detecting duplicates by counting the number of character substitutions needed to transform one string into another, and primarily targets names. Since user queries almost exclusively contain short keywords, the choice of this algorithm makes the most sense.

In order to make sure that the quality of the clusters was good, we only kept the results that satisfied the threshold of 0.8. The algorithm has been useful to deal with typographic errors frequently present in place names and to cluster search queries like *brusel* with their correct Dutch form, *brussel*. Perhaps more impressively, it also matched *Deinze O.L.V.* and *Onze Lieve Vrouwkerk Deinze*, two different query strings aiming to find the one and only *Onze Lieve Vrouw* (“Our Lady”) church in the Belgian city of Deinze. However, it should be noted that while trying to cluster different strings, one

does induce noise and it is nearly impossible to have an algorithm that matches different-but-similar words without false positives, such as “shirt” and “t-shirt”.

5 Results

First of all, results obtained while applying our 5-step method to our case study led to a significant reduction in the data to be analysed. At each stage (table 1), we were able to use filters and different processes to keep only the chosen data and to regroup very similar queries. To sum up, we moved from an initial data set consisting of more than 13 GB of data to a final file containing about 22 000 queried terms.

Table 1: Results obtained in terms of quantity¹⁵

Collecting	Parsing	Grouping	Cleaning	Clustering
Filtering 30 MB of “relevant” URLs from 13 GB of raw data	Extraction of 189,000 user queries from the URLs	Skimming up to 50,000 queries by keeping each one max 1×/visit	Harmonization from 50,000 to 37,000 different queries	Obtainment of 22,000 different queries

In terms of content, it is interesting to consider how the queries with the higher frequency count are affected by the method applied before retrieving them. There is no big surprise among the three most popular queries (table 2): there are either about parish registers (“parochieregisters” in Dutch and “registres paroissiaux” in French) or civil status registers (“etat civil” in French).

Table 2: The 3 most requested terms

After Parsing	After Grouping	After Normalising	After Clustering
parochieregisters	parochieregisters	etat civil	parochieregisters
etat+civil	etat civil	parochieregisters	etat civil
etat civil	registres paroissiaux	registres paroissiaux	registres paroissiaux

¹⁵ For clarification purposes, numbers have been rounded to the nearest unit for the “collecting” step and to the nearest thousand for all other steps.

More interesting results appear if we look at the 10 most popular queries. Thus, among the queries parsed from the raw data, three of them (“liege”, “tournai” and “ellezelles”, which are three Belgian place names) disappear within the next steps, being apparently less representative than expected. Similarly, among the top 10 queries resulting from clustering, new entities appear, which were absolutely not present in the previous steps, such as “burgerlijke stand” (civil status, in Dutch this time), “mariage” (wedding in French) or “acte de naissance” (birth certificates in French).

Eventually, in terms of contents itself, a more in-depth analysis should be carried out in collaboration with the staff members of the institution. However, it is already possible to notice that among the most requested terms, we find some global types of archive, such as parish registers, i.e. archive that were produced in the context of a parish. This finding reveals that current access to this type of archive might not be sufficiently visible, intuitive or effective.

6 Discussion and future work

Based upon the findings presented and analysed in this paper, it appears that text mining has the potential to help cultural heritage institutions to deal with huge data sets and obtain more representative results. However, the method presented here remains relatively labor-intensive and cannot be completely automated. Elements such as URL structures are variable and require human supervision to make the necessary adaptations. Such a process seems to be worthwhile as long as time and effort devoted to that analysis are considered as an initial investment which can lead to substantial savings for the future.

While first insights, based upon a narrow selection of the most requested terms within the digital catalogue, can be useful for the staff members of the Royal Archives of Belgium, they should not detain us from also considering the other results. As underlined by Khoo et al. (2008), “websites, as Internet nodes, exhibit many of the power law distributions typical of the Internet, characterised by a small number of data with high frequency counts at one end of the distribution, and a large number of data with low frequency counts at the other end.” Thus, considering only data with high frequency counts, such as the 10 most used queries, could overshadow other relevant data.

Future research will build on the method presented here in various ways. One possible way is to make comparisons in order to evaluate whether it is better to analyse queries as a whole, as performed here, or by parsing each word if they are composed of several terms. Also, more insights could be gained by identifying which words are the most often associated to some others, such as locations accompanying parish registers. Moreover, it would probably lead to more significant and interesting data to use a list of stop words including determinants and prepositions. Another research avenue could be to focus on visualising the evolution of queries in time. Eventually, one could resort to using named-entity recognition, which aims at detecting interesting entities in text and assign them a type, such as person, location or product, and therefore provides insights as to what kind of queries the visitors are interested in.

Acknowledgements

The authors would like to extend their gratitude to Raphaël Hubain, Seth van Hooland, Florence Gillet, Martin Vanbrabant, Gerald Haesendonck, Jill Hungenaert, Emmylou Haffner and Mathias Coeckelbergs. The research underlying the results presented in this article was funded by the Belgian Science Policy Office in the context of contract number BR/153/A5/MADDLAIN.

References

- Ceccarelli, Diego, Sergiu Gordea, Claudio Lucchese, Franco Maria Nardini, and Gabriele Tolomei (2011): Improving Europeana Search Experience Using Query Logs. In: *Research and Advanced Technology for Digital Libraries: International Conference on Theory and Practice of Digital Libraries, TPDL 2011* (pp. 384–395). Berlin, Heidelberg: Springer.
- Chase, Darren, Elizabeth Trapasso, and Robert Tolliver (2016): The Perfect Storm: Examining User Experience and Conducting a Usability Test to Investigate a Disruptive Academic Library Web Site Redevelopment. In: *Journal of Web Librarianship*, 10 (1), 28–44. [doi:10.1080/19322909.2015.1124740](https://doi.org/10.1080/19322909.2015.1124740)
- Dijkshoorn, Chris, Lora Aroyo, Guus Schreiber, Jan Wielemaker, and Lizzy Jongma (2014): Using Linked Data to Diversify Search Results a Case Study in Cultural Heritage. In: K. Janowicz et al. (Eds.): *Knowledge Engineering and Knowledge*

- Management. 19th International Conference, EKAW 2014, Linköping, Sweden, Proceedings* (pp. 109–120).
- Goyvaerts, Jan, and Steven Levithan (2012): *Regular Expressions Cookbook*. O'Reilly Media.
- Grimes, Carrie, Diane Tang, and Daniel M. Russell (2007): Query Logs Alone are not Enough. In: *Proceedings of the Workshop on Query Log Analysis: Social and Technology Challenges at the 16th International World Wide Web Conference (WWW 2007), Banff, Canada, May 8–12, 2007*. <http://static.googleusercontent.com/media/research.google.com/en/pubs/archive/34431.pdf>
- Jaro, M. A. (1989): Advances in Record Linkage Methodology as Applied to the 1989 census of Tampa Florida. In: *Journal of the American Statistical Society*, 84 (406), 414–420.
- Kathuria, Ashish, Bernard J. Jansen, Carolyn Hafernik, and Amanda Spink (2010): Classifying the User Intent of Web Queries using K-means Clustering. In: *Internet Research*, 20 (5), 563–581.
- Khoo, Michael, Joe Pagano, Anne L. Washington, Mimi Recker, Bart Palmer, and Robert A. Donahue (2008): Using Web Metrics to Analyze Digital Libraries. In: *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries (JCDL '08), Pittsburgh, USA, June 16–20, 2008*, New York (pp. 375–384).
- Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schütze (2009): *Introduction to Informational Retrieval* (online edition). Cambridge University Press. <http://nlp.stanford.edu/IR-book/pdf/irbookprint.pdf>
- Showers, Ben (2015): *Library, Analytics and Metrics: Using data to drive decisions and services*. Facet Publishing.
- van Hooland, Seth, Françoise Vandooren, and Eva Méndez Rodríguez (2011): Opportunities and Risks for Libraries in Applying for European Funding. In: *The Electronic Library*, 29 (1), 90–104.
- Winkler, William E. (1990): String Comparitor Metrics and Enhanced Decision Roles in the Fellegi-Sunter Model of Record Linkage. In: *Proceedings of the Section on Survey Research Methods*. Washington, DC: American Statistical Association (pp. 354–359).
- Woods, Roberta (2010): From Federated Search to the Universal Search Solution. In: *The Serials Librarian*, 58 (1), 141–148.
- Zavalina, Oksana (2007): Collection-Level User Searches in Federated Digital Resource Environment. In: Andrew Grove (Ed.): *Joining research and practice: social computing and information science / ASIS&T 2007, October 19–24, Milwaukee, Wisc.* Proceedings of the ASIS&T Annual Meeting; Vol. 44, Issue 1. Silver Spring, Md.: American Society for Information Science and Technology. [doi:10.1002/meet.1450440225](https://doi.org/10.1002/meet.1450440225)

Zavalina, Oksana, and Elena V. Vassilieva (2014): Understanding the Information Needs of Large-Scale Library Users. In: *Library Resources & Technical Services*, 58 (2), 84–99.

Session 5:
Metrics and Altmetrics

A Bibliometric Framework to Identify and Delineate Subfields of Research on Tribological Wear

Part One: Can We Identify Fundamental Issues by Clusters of Similar Journals?

Edgar Schiebel¹, Davide Bianchi², András Vernes²

¹ AIT Austrian Institute of Technology GmbH, Austria
edgar.schiebel@ait.ac.at

² AC2T research GmbH, Austria
{bianchi, vernes}@ac2t.at

Abstract

This contribution is about the first part of a bibliometric framework to identify new fundamentals in friction and wear in the field of tribology research. The first part covers clustering and mapping of bibliographically coupled journals to delineate subfields of tribological wear. The main objective of the whole study was to identify new theory based approaches and fundamentals on research about wear. The whole framework uses indicators from bibliometric elements about journals and conferences, disciplines, authors' keywords and research fronts extracted from publications. We collected slightly more than 5000 publications of relevant literature for the year 2015. In this contribution we report on structuring of journals and conference proceedings. The cluster analysis and a spring based one dimensional mapping of bibliographically coupled journals delivered some delineated disciplinary subfields like materials, mechanical engineering, dentistry, arthroplasty and machining. Clusters of journals on tribological wear delivered applied research as well as some more fundamental issues. However the analysis did not result in specific clusters of journals about new fundamental research on wear.

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 192–202.

Keywords: bibliometric framework; delineation of subfields; science mapping; bibliographic coupling; journals; tribology; wear

1 Introduction

The delineation of subfields of a broader research field attracted growing attention by the bibliometric community in the last years. We used recent findings in bibliometrics to construct a bibliometric framework to map and delineate research fronts of friction and wear in tribology. The key questions were: What are subfields in research on friction and wear? How can we identify fundamentals and mechanisms of wear by bibliometric methods?

The bibliometric framework was constructed in a way that a comprehensive set of publications could be analyzed from different points of view. First we had to define a search strategy and download a set of relevant publications. The bibliometric part consisted of an analysis of journals and conferences, disciplines, authors' keywords and research fronts by bibliographically coupled publications.

To define a proper search strategy we used the definition and fundamentals in wear from basic publications and textbooks. Jost (1966) first introduced the term tribology in his report of the Committee of the British Department (Ministry) of Education and Science. The term has its origin in the Greek word "tribo": "I ribe". He defined "tribology" as "The science and technology of interacting surfaces in relative motion – and of associated subjects and practices". It is strongly combined with friction, wear and lubrication. Nowadays fundamentals in practice and science deal with primary wear modes: abrasive wear, scratching; adhesive wear, galling, scuffing; fretting and fretting corrosion; erosive wear, cavitation impact, electro – arcing; rolling contact fatigue, spalling, delamination and tribo – corrosion, just to give some examples, see also Wear of Materials (2016).

In this paper we present the first part of the analytic framework: the identification of subfields by bibliographic coupling of journals and proceedings papers.

2 Methodology and data

Building the bibliometric framework as a whole we focused on relational bibliometric approaches using co-occurrence analysis, bibliographic coupling, co-citation and co-authorships. Price (1965) first published a basic concept for networks of publications when linking them directly by common references. Firstly he also introduced the term research fronts for a cluster of similar, bibliographically coupled publications. This method was adopted and is intensively used in the scientometric community nowadays. Boyack and Klavans (2010) examined the performance and accuracy of bibliographic coupling, co-citation analysis and direct citation is the least accurate mapping approach so far. Bibliographic coupling slightly showed better results than co-citation analysis and much better results than direct citation. Schiebel (2012) published a method for bibliographic coupling using a spring model (Kopcsa & Schiebel, 1998) to visualize research fronts by three-dimensional local densities. This methodology will be used to map the different elements of the bibliometric framework.

As a first part of the bibliometric framework we present the analysis of clusters of bibliographically coupled journals. The first work to couple journals was published by Narin, Carpenter and Berlt (1972). They described the interrelationship of scientific journals by direct citations amongst 275 selected journals in mathematics, physics, chemistry, biochemistry and biology. The authors found clear boundaries between disciplinary fields and identified cross disciplinary journals between the different fields. Recently Zhang, Xiaomin and Lili (2016) used journal coupling to study the similarity of disciplinary subjects from the subject-classification system of Chinese library classification.

Our approach in this work is to identify sub fields of tribological wear research on the level of bibliographically coupled journals. We measured the similarity of journals and proceedings by the relative number of common references of publications per journal/proceedings. The similarity grew with the number of common references in relation to the total number of references (Jaccard index). We did not use direct citation because it forces a direct link between two journals and common references are a weaker condition for a link and express better the similarity of the content of the documents published in different journals. We used the BibTechMon Software for the two dimensional positioning with the spring model.

For the collection of documents we used a search strategy defined by a Boolean combination of keywords like wear, material removal, material transfer, surface damage, surface degradation and surface deterioration. The keyword combination was applied with the topic feature of the Web of ScienceTM Core Collection. Due to the huge number of publications (141,499) we reduced the set of publications to the year 2015 and selected only relevant disciplines (excluding disciplines like neurosciences or psychology or archology, etc.). Applying the search strategy we downloaded 5143 documents for further analysis. The data set included 1000 sources (journals, proceedings of conferences ...) and 136,704 different cited references. The journals for this analysis were extracted from the set of documents where the journals were given as a source. We did not use another criterion like disciplines to select the journals. This also means that we did not use all published documents of the journals but only those in the set of our term based search strategy. Detailed information about the search strategy is given in Schiebel (2016).

The most relevant keywords were extracted with a modified TFIDF:

$$\text{TFIDF}_i = t_{i,j} * \log(N/n_i)$$

with $t_{i,j}$ as the frequency i (number of publications) of authors' keywords in publications (always 1 per publication) of cluster j , N the number of all publications and n_i the number of publications with keyword i . This definition of TFIDF made it possible to rank keywords by relevance for the clustered sources. The keywords were used to describe the content of the documents published in the journals of each cluster and on the level of the cluster.

3 Results

The research on tribological wear is carried out in different disciplines. The journals with the highest number of publication are "Wear", "Tribology International" and the "International Journal of Advanced manufacturing". They are thematically different and cover a broad spectrum of tribological issues. We find thematically more specific journals for surface and coatings, manufacturing, design and materials but also for chemistry, geoscience, medical issues (arthroplasty and dentistry), polymer science and others.

Additionally to the number of publications the column “Number of References” in table 1 gives the total number of unique and not multiple cited references in publications of the journal from our data set. For example: our data set included 251 publications of the Journal “WEAR” with a total number of 6881 cited different references.

Table 1: Journals for research on wear (tribology) by number of publications and cited unique references (not citations) in the year 2015

Journal	Number of Publications	Number of References
WEAR	251	6881
TRIBOLOGY INTERNATIONAL	172	5442
INTERNATIONAL JOURNAL OF ADVANCED MANUFACTURING TECHNOLOGY	169	4082
SURFACE & COATINGS TECHNOLOGY	140	4571
PLOS ONE	104	4593
APPLIED SURFACE SCIENCE	98	3278
TRIBOLOGY TRANSACTIONS	77	2162
RSC ADVANCES	70	2677
MATERIALS & DESIGN	67	2400
TRIBOLOGY LETTERS	67	2258

The references that were cited in publications were used to measure the similarity of journals. Just to demonstrate the clustering of coupled journals table 2 lists the co-occurrence of common references of 5 selected journals. Comparing the number of references of table 1 with the diagonal values of table 2 we find the same values. The highest co-occurring number of common references can be found for the journals “Wear” and “Tribology International” with 711 common references. How to be expected the journal

“Surface & Coating Technology” shows a high similarity to the tribology related journals 4 and 5 but not to “Plos One” and not too much to “The International Journal of Advanced Manufacturing Technology”. This outcome is fully appropriate, because as we will see later Plos One covers a broad number of different disciplines.

In a next step we clustered and mapped journals using the cited references of publications to measure the similarity with the Jaccard index. We expected that the clustering and mapping exercise would deliver different subfields of research on tribological wear and additionally would guide to new fundamentals in research on wear

Table 2: Cross table of co-occurrence of common references on publications in 5 selected journals

ID	Source	1	2	3	4	5
1	INTERNATIONAL JOURNAL OF ADVANCED MANUFACTURING TECHNOLOGY	4082		68	58	155
2	PLOS ONE		4593		14	20
3	SURFACE & COATINGS TECHNOLOGY	68		4571	298	301
4	TRIBOLOGY INTERNATIONAL	58	14	298	5442	711
5	WEAR	155	20	301	711	6881

The results are shown in figure 1 and 2. Figure 1 shows a circular representation of a cluster dendrogram (for the method of circular dendrograms see for example: How to ... 2016).

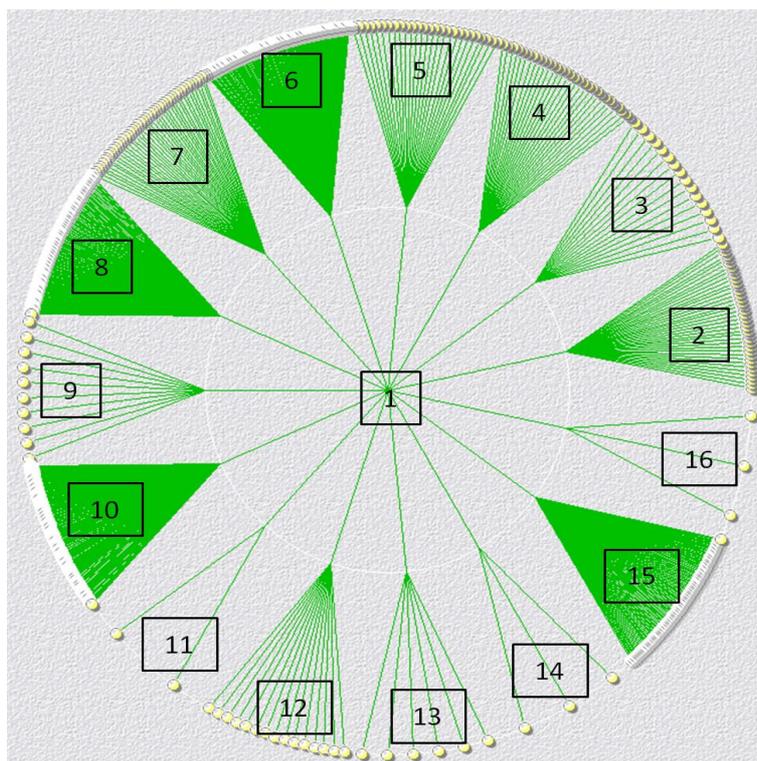


Fig. 1 Circular dendrogram of bibliographically coupled journals, Pearson agglomeration of Jaccard indexes, Ward linkage. The bubbles are the journals, “1” is the root and the numbers correspond with the identified clusters at the level of 15 selected clusters. Software: BibTechMon.

The clusters were calculated by the Pearson Correlation of the Jaccard index and the Ward agglomeration method. The dendrogram was cut at a level of easy to handle 15 clusters. This number was selected after trying different levels and numbers and reading the most relevant keywords ranked by the TFIDF. The selected clusters offered a high number of not “the same” keywords and a list of thematically consistent keywords per cluster. The name of a cluster was formally given by the journal with the highest number of publications.

Figure 2 shows a map of agglomerations of the journals calculated with a spring model. Journals in the same cluster of the cluster analysis were marked with the same color and the size of the bubbles represents the number of references cited in publications of each journal. Clusters of journals with the disciplines tribology, material science and mechanical engineering (2, 4, 5 and 6) are overlapping and located in the northern part of the map.

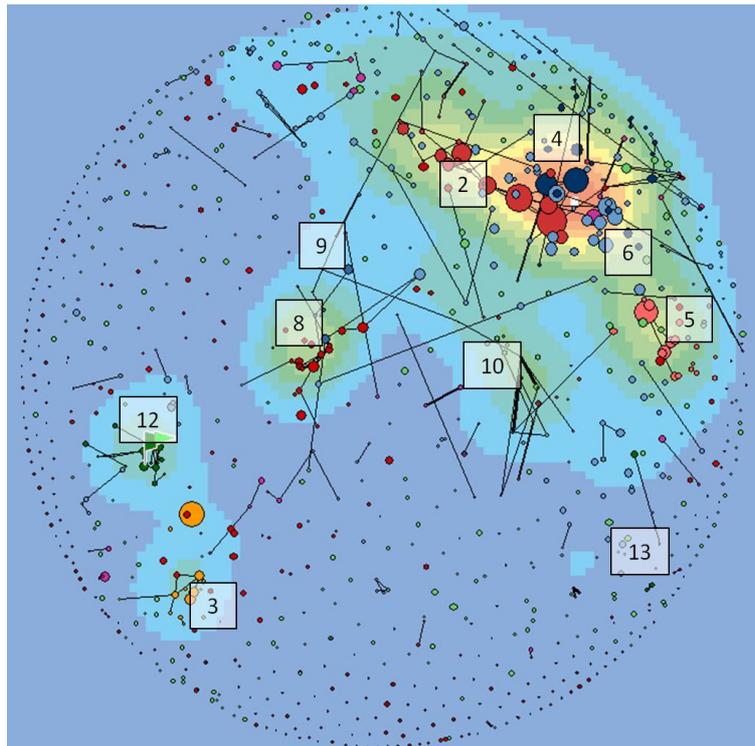


Fig. 2 2-dimensional mapping of bibliographically coupled journals, local distribution, positioned by a spring model, size: number of references, heat colors: local distribution of journals weighted by similarity. The numbers correspond with the cluster analysis and the identified clusters of journals. Software: BibTechMon.

Cluster 2 is formed by the journal “Wear”, by journals with “Tribology” in the title, materials and other mechanical engineering issues. We find publications on sliding wear, boundary lubrication and graphene as an emerging lubricant. Metal matrix composites together with research on wear resistance, micro-hardness and corrosion resistance are research issues of journals of Cluster 4.

Cluster 5 is a set of coupled journals related to manufacturing processes and tools. Research work in this cluster covers material removal, tool wear, surface roughness and cutting forces. The last cluster in this field is cluster 6. This is a huge cluster of 228 residual not bibliographically coupled journals from the materials and mechanical engineering area. The publications in this field deal with wear resistance tribology, mechanical property and corrosion what are issues of more general kind in tribology.

Clusters 3, 8 and 12 show a sharp delineation and are of different kind of disciplines than the main, big area mentioned above. Cluster 3 includes the journals Plos One, BMC Public Health, Medicine and Science in Sports and Exercise. It covers research on physical activity accelerometers and is not relevant for tribological wear because the documents matched with the term wear in the sense of people wearing accelerometers.

The cited references of the cluster “Plos One” couple additionally to tooth wear and dentistry that again fits with our field. This leads us to cluster 12 with journals like Archives of Oral Biology and Journal of Dentistry that cover publications on dental erosion, tooth wear and fluoride from medicine. Cluster 8 is formed by journals for total hip and total knee arthroplasty, biomechanics, prosthesis, polyethylene and biomaterials: Journal for Arthroplasty, Journal for Biomechanics and Acta Biomaterialia. Cluster 9 is close to cluster 8 and forms a link to the materials and engineering fields. It represents research on wear and lubrication of polymers and biomaterials, surfaces and interfaces and is covered by the following journals: Journal of Applied Polymer Science, Journal of Biomedical Materials Research Part B – Applied Biomaterials and Polymer Journal. Not mentioned clusters are spread over the landscape of figure 2

The journals of cluster 7 are spread over the thematic landscape and cover different issues on material research. With 281 Journals cluster 10 is the largest one and the journals are also spread over the thematic landscape. It does not represent a delineated field. Cluster 15 covers all remaining journals that are not coupled with each other by references. Cluster 11 and 14 are very small ones. Cluster 13 covers different journals: Mathematical Problems in

Engineering, Mechanism and Machine Theory as well as Minerals Engineering.

Table 2: Clusters of journals for research on wear (tribology), the name of each cluster is the journal with the highest number of references within the cluster, three most relevant authors' keywords measured by the TFIDF are listed in the last column

No.	Cluster Name (Number of Sources)	Keywords
2	WEAR (37)	sliding wear, wear testing, secondary lubrication
3	PLOS ONE (16)	physical activity, accelerometers
4	SURFACE & COATINGS TECHNOLOGY (29)	wear resistance, corrosion resistance, micro hardness
5	INTERNATIONAL JOURNAL OF ADVANCED MANUFACTURING TECHNOLOGY (26)	tool wear, surface roughness, cutting force
6	MATERIALS & DESIGN (228)	mechanical property, composite, corrosion
7	JOURNAL OF MATERIALS ENGINEERING AND PERFORMANCE (44)	corrosion and wear, Scanning Electron Microscopy (SEM), fatigue
8	JOURNAL OF ARTHROPLASTY (152)	total hip arthroplasty, polyethylene, total knee arthroplasty
9	JOURNAL OF APPLIED POLYMER SCIENCE (10)	wear and lubrication, polyolefins, degradation
10	JOURNAL OF MANUFACTURING SCIENCE AND ENGINEERING-TRANSACTIONS OF THE ASME (282)	absorption, nanocomposite, residual stress
11	MATERIALS CHEMISTRY AND PHYSICS (2)	composite material, nitrides, chemical synthesis
12	ARCHIVES OF ORAL BIOLOGY (15)	dental erosion, tooth wear, enamel
13	MATHEMATICAL PROBLEMS IN ENGINEERING (6)	Discrete Element Method (DEM), breakage, slider-crack mechanism
14	INTERMETALLICS (3)	metallic glass, solid-solution hardening, solid-solution hardening
15	INTERNATIONAL JOURNAL OF EARTH SCIENCES (147)	residual cluster, many different journals, no consistent keywords

4 Conclusions

In the first part of the application of a bibliometric framework to analyze subfields and research issues in tribological wear, we used the bibliographic coupling of journals that were identified by a set of documents collected with a string based search in Web of ScienceTM. The coupling was defined by common references in published documents on the level of the journals. We clustered the journals and calculated a two dimensional map where we positioned the journals due to their similarity defined by the Jaccard Index of the co-occurring references. The cluster analysis revealed some sharply delineated clusters in materials research, mechanical engineering, surfaces and composites on one hand. Wear is also subfield that consists of dental wear, arthroplasty, earth sciences and machining and tooling. In addition to the cluster analysis the two dimensional heat map of similar journals shows the compactness of the clusters and visualizes the content motivated relatedness of groups of journals. It was also used as a graphical retrieval tool to list keywords for each cluster.

We can conclude that also in research on wear Bradfords' law is valid and that the "core documents" would occur in journals with the highest number of publications. To easy identify new fundamentals in wear it would have been helpful to identify one or more clusters of journals that just cover such new fundamentals on wear. But it wasn't what a limitation of the presented approach is. The conclusion is that we have to look closer inside of the publications. Fundamental issues for wear should be found in Cluster 2 named by the journal "WEAR", in cluster 4 named by the journal "SURFACE & COATINGS TECHNOLOGY", cluster 6: "MATERIALS & DESIGN", cluster 11: "MATERIALS CHEMISTRY AND PHYSICS" and cluster 13: "MATHEMATICAL PROBLEMS IN ENGINEERING"

In future work we will compare the results to maps and clusters of disciplines, networks of authors and affiliations, research fronts based on bibliographically coupled publications as well as knowledge bases of co-cited references.

Acknowledgement

The work is part of an ongoing project called "Science Dynamics in Tribology – A Scouting Process with a Special Focus on Selected Research Issues" of the Austrian Center of Competence for Tribology. It is funded by the Austrian COMET Program managed by the FFG.

References

- Boyack, Kevin and Richard Klavans (2010): Co-Citation Analysis, Bibliographic Coupling, and Direct Citation: Which Citation Approach represents the Research Front Most Accurately? In: *Journal of the American Society for Information Science and Technology*, 61 (12), 2389–2404.
- How to plot a fan (Polar) Dendrogram in R? (2016, January 18). <http://stats.stackexchange.com/questions/4062/how-to-plot-a-fan-polar-dendrogram-in-r>
- Jost, Hans Peter (1990): Tribology – Origin and future. In: *Wear*, 136 (1), 1–17.
- Kopcsa, Alexander and Edgar Schiebel (1998): Science and Technology Mapping: A New Iteration Model for Representing Multidimensional Relationships. In: *Journal of the American Society for Information Science*, 49 (1), 7–17.
- Narin, Francis, Mark Carpenter and Nancy C. Berlt (1972): Interrelationships of scientific journals. In: *Journal of the American Society for Information Science*, 23, 323–331. [doi:10.1002/asi.4630230508](https://doi.org/10.1002/asi.4630230508)
- Price, Derek J. de Solla (1965): Networks of scientific papers. In: *Science*, 149, 510–515.
- Schiebel, Edgar (2012): Visualization of Research Fronts and Knowledge Bases by Three-Dimensional Areal Densities of Bibliographically Coupled Publications and Co-Citations. In: *Scientometrics*, 91 (2), 557–566. [doi:10.1007/s11192-012-0626-8](https://doi.org/10.1007/s11192-012-0626-8)
- Schiebel, Edgar (2016): Science Dynamics in Tribology – A Scouting Process with a Special Focus on Selected, Research Issues of the Austrian Center of Competence for Tribology Interim Report No. 1 to project AIT No. 1.61.00439.0.0, AIT-F&DP Reports, Vienna.
- Wear of materials (2016): Conference Program, Wear of Materials 2017, 21st International Conference on Wear of Materials, 26–30 March 2017 | Hilton Long Beach, California, USA. <http://www.wearofmaterialsconference.com/<10.4.2016>>
- Zhang, Jing, Liu Xiaomin and Wu Lili (2016): The study of subject-classification based on journal coupling and expert subject-classification system. In: *Scientometrics*, 107 (3), 1149–1170.

Are Altmetrics Effective in Transdisciplinary Research Fields?

Altmetric Coverage of Outputs in Educational Research

*Elisabeth Vogler, Christoph Schindler,
Alexander Botte, Marc Rittberger*

DIPF, Germany
vogler@dipf.de

Abstract

This paper analyzes the coverage of altmetrics for heterogeneous scientific and extra-scientific outputs in a transdisciplinary research field. The transdisciplinary field of educational research is used as a case study to get first insight how current altmetric tools cover the field on the levels of its general publication output, and on the level of research relevant journals. Additionally, an experimental approach analyzes the Twitter mentions of a transdisciplinary research report.

Keywords: altmetrics; transdisciplinarity; tweet analytics; educational research; publications practices

1 Introduction

In this paper, the coverage of altmetrics is analyzed in a transdisciplinary research field characterized by heterogeneous scientific and extra-scientific outputs. Recently a fundamental transformation of the scientific landscape has been discussed as transdisciplinary research and has become a major topic in science policy agendas. Transdisciplinary research can be described

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 203–215.

as a reflexive research approach which integrates societal problems into scientific issues, involves heterogeneous scientific and extra-scientific insights and contributes to both societal and scientific progress (cf. Jahn, Bergmann & Keil, 2012; Gibbons, 1994). Classical scientometrical approaches are mostly limited to an intra-scientific measurement. Altmetrics offer potentials to include extra-scientific communication and have been discussed in the broader context of scientific impact as societal impact (Barnes, 2015; Bornmann, 2014a, 2015; Holmberg et al., 2015). Furthermore, with respect to the heterogeneity of transdisciplinary research outputs, altmetrics are able to integrate a broader range of research products into the scientific evaluation apparatus like datasets, software, algorithms, grey literature, and slides (Bornmann, 2014a; Priem, 2014; Zahedi, Fenner & Costas, 2014). On the other hand, the measurement of non-scientific output is difficult (Koier & Horlings, 2015) and so far no systematic analysis of the accuracy of altmetrics in transdisciplinary research has been presented.

This paper uses the transdisciplinary field of educational research as a case study. Educational research is characterized as problem-oriented, disciplinarily heterogeneous (e.g. psychology, social and political sciences, economics) with a strong alignment to educational practice (Baumann, 2005; Dees, 2015; Deutscher Bildungsrat, 1974). The transdisciplinary character of educational research and its heterogeneous disciplinary publication cultures and practices are used in this study to identify challenges and boundaries of altmetrics, describe transnationality and discuss societal impact.

In this study different methodological ways are tested to get first insight how current altmetric tools cover the field on the levels of its general publication output, and on the level of research relevant journals. An additional experimental approach analyzes the Twitter mentions of a continuous research report on the situation of education in Germany (Autorengruppe Bildungsberichterstattung, 2016). The education report was chosen as a relevant transdisciplinary output in educational research (cf. Dees, 2015: 178), which addresses the extra-scientific world of educational politics, administration, and the general public.

2 State of the art

Transdisciplinary research is widely discussed and a variety of conceptualizations exists (Gibbons, 1994; Jahn, Bergmann & Keil, 2012; Nowotny, Scott & Gibbons, 2001). A common setting can be described for this study as: Integrating societal problems into scientific issues, involving heterogeneous scientific and extra-scientific insights and contributing to both societal and scientific progress (Jahn, Bergmann & Keil, 2012; Klein, 2008). Respectively, difficulties emerge concerning operationalization and measurement. The integration of heterogeneous disciplinary knowledge into research (interdisciplinarity) is regarded as difficult to measure in practice (Huutoniemi et al., 2010), for which accounts scientific output (mainly publications) (Wagner et al., 2011). The situation is aggravated by integrating the extra-scientific world (i.e. problems, knowledge, stakeholders, and progress). So, for example, stakeholders and scientists differ in their agendas of interests, which limits the usage of bibliometric methods based on scientific publication data bases (Koier & Horlings, 2015). Furthermore, some authors have argued that the intra-scientific alignment of the publication data bases (i.e. WOS) allows no statements about societal impact (cf. Holmberg et al., 2015). Others emphasize the relational aspects of transnationality, like societal progress, which are difficult to grasp with metrics (Klein, 2008; Koier & Horlings, 2015).

For some time altmetrics have been discussed as so called alternative metrics promising the extension of current citation-based impact factors (Priem et al., 2010). Various potentials are articulated concerning transdisciplinary research (Koier & Horlings, 2015), especially for addressing societal progress (Bornmann, 2014b, 2014a). Altmetrics involves the extra-scientific world (knowledge, stakeholders) and extends the range of research products for scientific evaluation (Bornmann, 2014a, 2015; Priem, 2014; Zahedi, Fenner & Costas, 2014). A further advantage is seen in the more timely response of Twitter citations in comparison to citations in publications (Bornmann, 2016). On the other hand, a series of related limitations have been articulated in altmetrics ranging from coverage, consistency and traceability of data sources and aggregators (Chamberlain, 2013; Zahedi, Fenner & Costas, 2014, 2015) to the dependency on communication and publication practices of different scientific communities (Costas, Zahedi & Wouters, 2014; Peters et al., 2014; Zahedi, Costas & Wouters, 2014).

Current research in altmetrics still preferably deals with questions, if altmetrics reach already a representative level of coverage and significance in comparison to traditional bibliometrics. Most of the altmetric studies are focused on science and medicine disciplines (e.g. Andersen & Haustein 2015, Bar-Ilan 2014, Adams & Loach 2015). Recently there were some multidisciplinary studies covering also social sciences and arts and humanities. Mohammadi and Thelwall (2014) analyzed Mendeley readers' data of WoS articles from 2008 from different disciplines. With 39% readership statistics articles in the transdisciplinary field of the educational and educational research represented a quite high readership quotation, whereas the humanities articles showed only 13% of Mendeley readers.

Including the whole range of social media sources Costas et al. (2014) found altmetrics data for 22.5% of the arts and humanities articles. In 2015 they published a multi-disciplinary study which analyzed the "thematic orientation of publications mentioned on social media" (Costas, Zahedi & Wouters, 2015). Summing-up the authors stated: "The humanities, natural sciences, and engineering disciplines have a much lower presence of social media metrics. Twitter has a stronger focus on general medicine and social sciences. Other sources (blog, Facebook, Google+, and news media mentions) are more prominent in regards to multidisciplinary journals" (ibid.: 260). Peters et al. (2014) explored the altmetric mentions for publications (journal articles and book chapters) by twelve Leibniz-Institutes (2011–2012). An institute from the Humanities and one from the educational sciences were selected (section A). Both institutes from section A had with 32% and 30% in comparison to the other sections the least mentions in social media.

The heterogeneity of publication practices has been addressed in social science and humanities in general (Fry & Talja, 2004; Nederhof, 2006) and in particular in educational research (Singleton et al., 2015; Dees, 2008), where authors publish aside from journals mainly in monographs and compilations. In transdisciplinary research fields this heterogeneity needs to be considered because various research communities and others are involved. In educational research the heterogeneity of publication practices affects the significance of the publication format 'article'. Articles, the favored publication format of citation data bases, are significantly more used in psychological (36.2%) oriented educational research than in sociological (27.4%) and genuine educational (28.9%) research (Singleton et al., 2015: 83). Thus, research communities with a low article orientation are misrepresented in arti-

cle based citation and reference databases like Web of Science (WOS) and Scopus, which are used as a reference for altmetrics.

3 Data and methodology

The survey on the coverage of publication practices in educational research was carried out in subject-specific databases that as far as possible portray the disciplinary heterogeneity. For the field coverage level of publications (as of June 2016) in educational research, the German Education Index (GEI) was used (2010–2015), which offers a broad spectrum of different publication types (e.g. monographs) including grey literature and practice-oriented publications outside of citation-based scientific ranking systems. For all GEI articles with the identifier DOI the altmetrics for the publications were retrieved via the Altmetric.com-API using R and the package rAltmetric on 10 May 2016.

For the analysis of educational journals, the index of editing characteristics of educational research journals DEPOT was used, a proved collection of periodicals relevant for educational research in Germany. All ISSNs of the journals indexed as of 26 February 2016 were inserted in Altmetric.com via the Altmetric.com explorer to retrieve all journal articles indexed in the Altmetric.com database. A further collection of Twitter-mentions addresses a continuous research report on the situation of education in Germany (Autorengruppe Bildungsberichterstattung, 2016). The mentions were recorded via the Twitter-Search API with NodeXL. The tweets were collected for seven days from the day of publication (June 16, 2016).

4 Coverage of outputs in educational research

4.1 Coverage of educational research publications

The coverage of educational research publications based on the German Education Index (2010–2015) reveals that 21.2% of the publications with DOI have altmetric data. 17.5% of these publications with DOI were mentioned at

Twitter, followed by Blogs (2.3%), Google+ (1.0%), and News (0.2%). The high share of Twitter mentions is in line with the results of previous studies (e.g. Costas, 2015). But only 12% of the collected publication output disposed of a DOI (16,076 out of 134,301) and could be mapped with altmetrics. Concerning the full range of collected publications (2010–2015) and the circumstances of aggregation just 2.5% ($n = 3,404$) could be identified and connected to altmetric data (table 1). A missing DOI is apparently a profound coverage limitation of altmetric studies.

Table 1: Altmetrics for the publications in the German Education Index (2010–2015)

Year	All Pub.	Pub. with DOI	Pub. with Altmetrics (PubA/DOI) (PubA/GEI)	Twitter	Face-book	Blogs	News	Google+	Wiki-pedia
2010 to 2015	134,301	16,076	3,404 (21.2%) (2.53%)	2,816 (17.5%) (2.1%)	541 (3.4%) (0.4%)	368 (2.3%) (0.27%)	29 (0.2%) (0.02%)	162 (1.0%) (0.12%)	69 (0.43%) (0.05%)

A more precise picture emerges when considering the different types of publication used in a transdisciplinary research field with its heterogeneous publication practices. Figure 1 shows the distribution of publication types in the educational database GEI (2010–2015).

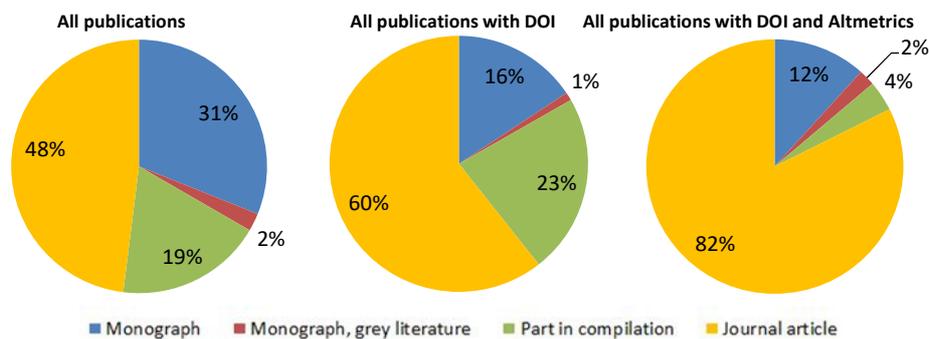


Fig. 1 Proportion of publication types in the GEI (2010–2015)

Looking at the relation between publication type and availability of DOIs we find: The share of journal articles in the database increases from 48% of all publications to 60% of the publications with DOI and to 82% of the publications with altmetric data. Other major publication types in the field like

monographs and parts of compilations decrease from 52% (all) to 40% (with DOI) to 18% (altmetrics). Taking into account that different research communities in transdisciplinary fields follow different publication practices in terms of preferred publication types, it seems probable that non-journal based fields of educational research face disadvantages when altmetric measurement is applied.

While the detection of impact of research on society is seen as a great potential for altmetrics, the coverage of its output and the capturing of feedback from multiple groups of audience also play a central role in a transdisciplinary field. The study also analyzed the role of language on impact in social media. While publications in German are at 72% in the GEI corpus and thus predominant, German publication output with DOI is 43% and only 5% are covered with altmetric data (fig. 2).

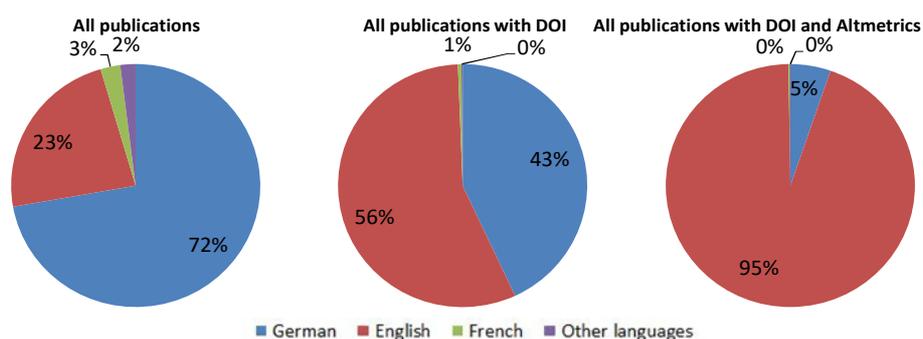


Fig. 2

Proportion of publication language in the German Education Index (2010–2015)

Given these conditions, it seems more difficult to address non-English target groups. Against the background of previous studies which focus nearly completely on the predominantly English publications of the WoS more research seems to be needed to reveal the influence of language on the coverage of altmetrics.

4.2 Coverage of educational research journals

The index of editing characteristics of educational research journals DEPOT (as of February 2016) describes and qualifies 310 journals building a solid base for analyzing the coverage of altmetrics on a journal level in educational research, but with a focus on German journals (94%). A query of ISSNs on

25 February 2016 revealed that 1,952 articles have altmetric data, sourced from 27 journals, i.e. 8.7% of all DEPOT journals. Reduced to the years 2010–2015 and after revision 1,427 articles from 23 journals have altmetric data, as described in table 2. 12 (52%) of these journals are English language journals, 11 (48%) are German. Due to the data collection method, direct via ISSN from Altmetric.com, it is not possible to have an overview of all publications in DEPOT journals or the publications with DOI. The output was a list of all articles with altmetric mentions.

Table 2: Altmetrics of Educational Research Journals (DEPOT 2010–2015)

Years	All articles with altmetric data	Twitter	Face-book	Blogs	News	Google+	Wiki-pedia
2010 to 2015	1,427 (100%)	1,323 (92.7%)	259 (18.2%)	74 (5.2%)	53 (3.7%)	160 (11.2%)	21 (1.5%)

18 of these journals have less than 50 articles with mentions covered by Altmetric.com and 9 journals have less than 10 mentioned articles. Considering the previously described language variety and its constraints for addressing societal progress, only three of the first 10 journals are German journals.

4.3 Coverage of Twitter-mentions for a transdisciplinary report

On 16 June 2016 the report ‘Bildung in Deutschland 2016’ (Autorengruppe Bildungsberichterstattung, 2016) was published. The report is an indicator-based study of the German educational system as a whole. This report is a scientific output, compiled by a transdisciplinary team of educational researchers; its targets are educational policy, educational administration and the broader public. Therefore, the report is an exemplary publication output in the transdisciplinary area of educational research.

The established way to search for altmetric mentions by querying an identifier returned no results. Based on the Twitter-Search-API and a query for ‘bildungsbericht’ there were 405 unique tweets for the time period from June 16 to June 22, 2016. In addition to the tweets there were 447 retweets from 134 unique tweets. 386 tweets contain URLs to 85 different domains. The most referenced domain is faz.net (website of Frankfurter Allgemeine Zeitung) with 52 links, followed by bildungsbericht.de (the official report website) with 38 links (cf. table 3). Among the ten most mentioned domains there

are seven links to established mainstream news sites. This points to an extra-scientific attention of the report and indicates an impact on a discourse level. Already (Costas, 2015) found out that multidisciplinary topics tend to draw the attention of discourse oriented media like news and blogs

Table 3: The ten most frequently mentioned domains

Domain	No. of tweets	Category of domain
faz.net	52	Newspaper
bildungsbericht.de	38	Website of the education report
spiegel.de	35	News magazine
tagesschau.de	17	Television news service
sueddeutsche.de	16	Newspaper
feedburner.com	14	Web feed management provider
zeit.de	14	Newspaper
bildungsserver.de	13	Information portal for education
deutschlandfunk.de	12	Broadcasting radio station
tagesspiegel.de	9	Newspaper

At the day of the publication of the report there was a big blip in the amount of mentions that dropped the following days. 65.5% (558) of the collected tweets were created on June 16, 2016; 14.1% (120) tweets were created the day after the publication on June 17, 2016. 640 persons were identified as active, 302 of whom sent tweets and 361 sent retweets. 43% (302) of the active persons just posted tweets, 53% (361) only retweeted, and 4% (23) did both. These results confirm altmetrics as an extremely real-time indicator of communication about publications evoking high societal attention. But the attention of societal groups may also be rather short-winded, whereas scholarly resonance will follow later, and it is still questionable whether altmetrics will cover intra-scientific resonance (citations) sufficiently.

5 Discussion

In this study, we analyzed the coverage of altmetrics with respect to heterogeneous publication practices and outputs in a transdisciplinary field, namely educational research. Instead of high percentages (39%) of altmetric

coverage like at WoS articles with DOI matchings with Mendeley readers data (Mohammadi & Thelwall, 2014) or 30% of altmetric data for articles with DOI of an educational research institute (Peters et al., 2014), 21% of all publications with DOI in the German Education Index (2010–2015) and therefore just 2.5% of all publications in the German Education Index (2010–2015) have measurable altmetric data. The fact that the German database GEI comprises a high majority of German language publications and a relatively low number of journal articles with DOI obviously diminishes the share of altmetric data.

On one hand, this indicates that the population and the specific publication culture of a discipline (here educational research in Germany) has to be considered in altmetrics studies. On the other hand, it shows the need that altmetric aggregations like Altmetric.com have to extend the restricted range of the single identifier DOI. For the time being the collection and aggregation process of altmetrics excludes major parts of the heterogeneity of publication practices in a transdisciplinary research field. Further, multilingual and book oriented publication cultures are disadvantaged by altmetrics' predominance of journal articles and English language publications. Currently altmetric studies deal with similar problems as bibliometric studies in terms of data collection and their limitation to journal articles. A central goal of altmetrics, i.e. addressing practice-oriented areas and locally oriented publications (cf. Koier & Horlings, 2015), is thus constrained.

Nevertheless, the study indicates potentials of social media analytics for transdisciplinary research outputs, exceeding traditional bibliometric. On the basis of an extended future set of measurable identifiers the communication of transdisciplinary publications will be covered much better and allow a monitoring of a much broader attention. The example of the report for education shows that extra-scientific discourse on research outputs referenced by mainstream media and beyond can be made visible and assessable.

A range of circumstances give reason to anticipate positive developments for more descriptive approaches in prompt monitoring of discourses about transdisciplinary research fields and outputs. In longer time periods trend identification or contextual enrichments of research outputs can be implemented in research infrastructures (e.g. reference databases). Coverage of altmetric instruments is still growing, altmetrics aggregators like Altmetrics.com involve further identifiers like the URN and short messages are increasingly being used at international conferences. Altmetrics thus face big

challenges but also bear potential for analyzing scientific outputs beyond scientific impact.

References

- Adams, Jonathan and Tamar Loach (2015): *Altmetric mentions and the communication of medical research. Disseminating research outcomes outside academia: Digital Research Reports*. London: Digital Science.
- Andersen, Jens Peter and Stefanie Haustein (2015): Influence of Study Type on Twitter Activity for Medical Research Papers. In: A. A. Salah (Ed.): *Proceedings of ISSI 2015, Istanbul*. Istanbul: Bogaziçi.
- Autorengruppe Bildungsberichterstattung (Ed.) (2016): *Bildung in Deutschland 2016*. Bielefeld: Bertelsmann, W.
- Bar-Ilan, Judit (2014): Astrophysics publications on arXiv, Scopus and Mendeley: a case study. In: *Scientometrics*, 100 (1), 217–225. [doi:10.1007/s11192-013-1215-1](https://doi.org/10.1007/s11192-013-1215-1)
- Barnes, C. (2015): The Use of Altmetrics as a Tool for Measuring Research Impact. In: *Australian Academic & Research Libraries*, 46 (2), 121–134.
- Baumann, U. (2005): Interdisziplinarität, Transdisziplinarität, Integralität. In: *Tertium comparationis*, 11 (1), 3–14.
- Bornmann, L. (2014a): Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. In: *Journal of Informetrics*, 8 (4), 895–903.
- Bornmann, L. (2014b): Validity of altmetrics data for measuring societal impact: A study using data from Altmetric and F1000Prime. In: *Journal of Informetrics*, 8 (4), 935–950.
- Bornmann, L. (2015): Usefulness of altmetrics for measuring the broader impact of research. In: *ASLIB Journal of Information Management*, 67 (3), 305–319.
- Bornmann, L. (2016): Measuring impact in research evaluations: A thorough discussion of methods for, effects of and problems with impact measurements. In: *Higher Education*. [doi:10.1007/s10734-016-9995-x](https://doi.org/10.1007/s10734-016-9995-x)
- Chamberlain, S. (2013): Consuming Article-Level Metrics: Observations and Lessons. In: *Information Standards Quarterly*, 25 (2), 4–13.
- Costas, R., Z. Zahedi, P. Wouters (2014): Do altmetrics correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary

- perspective. In: *Journal of the Association for Information Science and Technology*, 66 (10), 2003–2019. [doi:10.1002/asi.23309](https://doi.org/10.1002/asi.23309)
- Costas, R., Z. Zahedi, P. Wouters (2015): The thematic orientation of publications mentioned on social media. In: *ASLIB Journal of Information Management*, 67 (3), 260–288.
- Dees, W. (2008): Innovative Scientometric Methods for a Continuous Monitoring of Research Activities in Educational Science. In: *Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting*. Berlin.
- Dees, W. (2015): *Bildungsforschung – eine bibliometrische Analyse des Forschungsfeldes*. PhD thesis, Humboldt-Universität zu Berlin. <http://edoc.hu-berlin.de/dissertationen/dees-werner-2014-11-28/PDF/dees.pdf>
- Deutscher Bildungsrat (1974): *Aspekte für die Planung der Bildungsforschung*. Klett.
- Fry, J., and S. Talja (2004): The cultural shaping of scholarly communication. In: *Proceedings of the American Society for Information Science and Technology*, 41 (1), 20–30.
- Gibbons, M. (1994): *The new production of knowledge: The dynamics of science and research in contemporary societies*. Los Angeles: Sage.
- Holmberg, K., F. Didegah, T. Bowman, and T. Kortelainen, T. (2015): Measuring the societal impact of open science—Presentation of a research project. In: *Informaatio tutkimus*, 34 (4).
- Huutoniemi, K., J. T. Klein, H. Bruun, and J. Hukkinen (2010): Analyzing interdisciplinarity: Typology and indicators. In: *Research Policy*, 39 (1), 79–88. [doi:10.1016/j.respol.2009.09.011](https://doi.org/10.1016/j.respol.2009.09.011)
- Jahn, T., M. Bergmann, and F. Keil (2012): Transdisciplinarity: Between mainstreaming and marginalization. In: *Ecological Economics*, 79, 1–10.
- Klein, J. T. (2008): Evaluation of interdisciplinary and transdisciplinary research: a literature review. In: *American journal of preventive medicine*, 35 (2), S116–S123.
- Koier, E., and E. Horlings (2015): How accurately does output reflect the nature and design of transdisciplinary research programmes? In: *Research Evaluation*, 24 (1), 37–50. [doi:10.1093/reseval/rvu027](https://doi.org/10.1093/reseval/rvu027)
- Mohammadi, E., and M. Thelwall (2014): Mendeley readership altmetrics for the social sciences and humanities. In: *Journal of the Association for Information Science and Technology*, 65 (8), 1627–1638.
- Nederhof, A. J. (2006): Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review. In: *Scientometrics*, 66 (1), 81–100.
- Nowotny, H., P. Scott, and M. Gibbons (2001): *Re-thinking science: Knowledge and the public in an age of uncertainty*. Cambridge, UK: Polity.

- Peters, I., A. Jobmann, A. Eppelin, C.P. Hoffmann, S. Künne, S., and G. Wollnik-Korn (2014): Altmetrics for large, multidisciplinary research groups: a case study of the Leibniz Association. In: *Assessing libraries and library users and use*. Part II. Zadar.
- Priem, J. (2014): Altmetrics. In: B. Cronin and C. R. Sugimoto (Eds.): *Beyond bibliometrics. Harnessing multidimensional indicators of scholarly impact* (pp. 263–287). Cambridge, Mass.: The MIT Press.
- Priem, J., D. Taraborelli, P. Groth, and C. Neylon (2010): altmetrics: a manifesto. <http://www.altmetrics.org/manifesto>.
- Singleton, K., V. Kuhberg-Lasson, U. Sondergeld, and J. Schultheiß (2015): Publikationen der Bildungsforschung. In: A. Botte et al. (Eds.): *Monitoring Bildungsforschung*. Bad Heilbrunn: Klinkhardt, Julius (pp. 67–106).
- Wagner, C. S., J. D. Roessner, K. Bobb, J. T. Klein, K. W. Boyack, J. Keyton, (2011): Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. In: *Journal of Informetrics*, 5 (1), 14–26.
- Zahedi, Z., R. Costas, and P. Wouters (2014): How well developed are altmetrics? A cross-disciplinary analysis of the presence of ‘alternative metrics’ in scientific publications. In: *Scientometrics*, 101 (2), 1491–1513.
- Zahedi, Z., Fenner, M., and R. Costas (2014): How consistent are altmetrics providers? Study of 1000 PLOS ONE publications using the PLOS ALM, Mendeley and Altmetric.com APIs. <https://doi.org/10.6084/m9.figshare.1041821.v2>
- Zahedi, Z., M. Fenner, and R. Costas (2015): Consistency among altmetrics data provider/aggregators: what are the challenges? In: *altmetrics15. the 2015 Altmetrics Workshop, Amsterdam, 9 October 2015*.

Session 6:
(Social) Media Analysis

news-please

A Generic News Crawler and Extractor

*Felix Hamborg, Norman Meuschke,
Corinna Breitingner, Bela Gipp*

Department of Computer and Information Science
University of Konstanz, Germany
[firstname.lastname]@uni-konstanz.de

Abstract

The amount of news published and read online has increased tremendously in recent years, making news data an interesting resource for many research disciplines, such as the social sciences and linguistics. However, large scale collection of news data is cumbersome due to a lack of generic tools for crawling and extracting such data. We present news-please, a generic, multi-language, open-source crawler and extractor for news that works out-of-the-box for a large variety of news websites. Our system allows crawling arbitrary news websites and extracting the major elements of news articles on those websites, i.e., title, lead paragraph, main content, publication date, author, and main image. Compared to existing tools, news-please features full website extraction requiring only the root URL.

Keywords: news crawler; news extractor; scraper; information extraction

1 Introduction and background

News articles are an interesting data source for researchers in many domains. For instance, social scientists heavily rely on news for performing framing analyses, i.e. studying how people interpret situations and activities. News

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 218–223.

are a well-suited data source for this kind of analysis, since they reflect which events received public attention and how these events were portrayed. Some news data sets, such as RCV1 (Lewis et al., 2004), are freely available. However, researchers often need to compile their own dataset, e.g., to include news published by specific outlets or in a certain time frame. Due to the lack of a publicly available, integrated crawler and extractor for news, researchers often implement such tools redundantly. The process of gathering news data typically consists of two phases: (1) *crawling news websites* and (2) *extracting information from news articles*.

Crawling news websites can be achieved using many web crawling frameworks, such as *scrapy* for Python (Kouzis-Loukas, 2016). Such frameworks traverse the links of websites, hence need to be tailored to the specific use case.

Extracting information from news articles is required to convert the raw data that the crawler retrieves into a format that is suitable for further analysis tasks, such as natural language processing. Information to be extracted typically includes the headline, authors, and main text. *Website-specific extractors*, such as used in (Meschenmoser et al., 2016; Paliouras et al., 2008), must be tailored to the individual websites of interest. These systems typically achieve high precision and recall for their extraction task, but require significant initial setup effort in order to customize the extractors to a set of specific news websites. Such website-specific extractors are most suitable when high data quality is essential, but the number of different websites to process is low.

Generic extractors are intended to obtain information from different websites without the need for adaption. They use heuristics, such as link density and word count, to identify the information to be extracted. Our literature review and experiments have shown that *Newspaper* (Ou-Yang, 2013) is currently one of the most sophisticated and best performing news extractors. It features robust extraction of all major news article elements and supports more than ten languages. *Newspaper* includes basic crawling, but lacks full website extraction, auto-extraction of new articles, and news content verification, i.e. determining whether a page contains a news article. The extraction performance of other frameworks, such as *boilerpipe* (Kohlschütter, Fankhauser & Nejd, 2010), *Goose* (Labs, 2016), and *readability* (Baburov, 2010) is lower than that of the *Newspaper* tool. Furthermore, these latter tools do not offer support for crawling websites.

To our knowledge, no available tool fully covers both the crawling and extraction phase for news data. Web crawler frameworks require use-case specific adaptations. News extractors lack comprehensive crawling functionality. Existing systems lack several key features, particularly the capability (1) to extract information from *all* articles published by a news outlet (*full website extraction*) and (2) to auto-extract newly published articles. With *news-please*, we provide a system that addresses these two weaknesses using a generic crawling and extraction approach. The following section presents *news-please* in detail.

2 System overview

news-please is an open-source news crawler and extractor written in Python developed to meet five requirements: (1) broad coverage – extract news from any outlet’s website, (2) full website extraction, (3), high quality of extracted information, (4) ease of use – simple initial configuration, and (5) maintainability. Where possible, *news-please* uses existing state-of-the-art tools, which we extended with functionality to meet the outlined requirements. This section describes the processing pipeline of *news-please* as shown in figure 1.

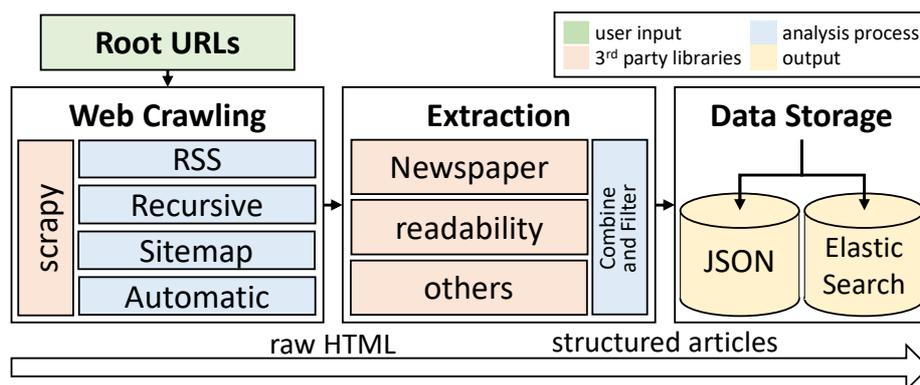


Fig. 1 Pipeline for news crawling and extraction

Root URLs. Users provide URLs that point to the *root* of news outlets' websites, e.g., nyt.com. For each root URL, the following tasks are performed.

Web Crawling. news-please performs two sub-tasks in this phase. (1) The *crawler* downloads articles' HTML, using the *scrapy* framework. (2) To find all articles published by the news outlet, the system supports four techniques: *RSS* (analyzing RSS feeds for recent articles), *recursive* (following internal links in crawled pages), *sitemap* (analyzing sitemaps for links to all articles), and *automatic* (tries sitemaps, falls back to recursive in the case of an error). The approaches can also be combined, e.g., by starting two news-please instances in parallel, one in automatic mode to get all articles published so far, and another instance in RSS mode to retrieve recent articles.

Extraction. We use existing state-of-the-art news extractors to obtain the desired information, i.e., title, lead paragraph, main content, author, date, and main image. In preliminary tests (cf. sect. 3), we evaluated the performance of four extractors (*boilerpipe*, *Goose*, *Newspaper*, and *readability*). *Newspaper* performed best for all news elements combined followed by *readability*. Thus, we integrated both extractors into *news-please*. Because both *Newspaper* and *readability* performed poorly for extracting publication dates, we added a regex-based date extractor (Geva, 2016). Our component-based design allows easily adding or removing extractors in the future. Currently, news-please combines the results of the extractors using rule-based heuristics. We discard pages that are likely not articles using heuristics, such as link-to-headline ratio, and metadata filters.

Data Storage. news-please currently supports writing the extracted data to JSON files and to an Elasticsearch interface.

3 Extraction performance

In tests, we compared the extraction performance of *news-please* against the four extractors: *boilerpipe*, *Goose*, *Newspaper*, and *readability*. In total, we selected 20 articles from 20 news websites (the top 15 news outlets by global circulation and five major outlets in Germany) and manually assessed the quality of extracted information using a four-point-scale: (A) perfect; (B) good: the beginning of an element is extracted correctly, later information is

missing or information from other elements is added, (C) poor: in addition to (B), the beginning of an element is not extracted entirely correctly, (D) unusable: much information is missing or from other elements. *news-please* performed particularly well for titles (82% in category A or B), description (76% in A or B), date (70% in A or B), and main image (76% in A or B). For other elements, the extraction quality can still be improved: main content (62% in A or B) and author (34% in A or B). Overall, *news-please* performed better than the included extractors individually.

4 Conclusion and future work

We present *news-please*, the first integrated crawler and information extractor specifically designed for news articles. *news-please* is able to crawl *all* articles of a news outlet including articles published during the crawling process. The system combines the results of three state-of-the-art extractors. For high maintainability and extendibility, *news-please* allows inclusion of additional extractors and adaption to use-case-specific requirements, e.g., by adding a SQL result writer. In tests, we found that *news-please* achieves a higher extraction quality than the individual extractors. By integrating both the crawling and extraction task, researchers can gather news faster and with less initial effort and long-term effort. In the future, we will focus on improving the extraction of an article's main content, supporting more languages, improving the combination of extracted content elements, and evaluating the performance of *news-please* in more detail. The code has been made available under an Apache 2 license at: <https://github.com/fhamborg/news-please>.

Acknowledgements

This work has been supported by the Carl Zeiss Foundation. We also thank the students at the University of Konstanz who helped realize *news-please*: M. Bock, M. Fried, J. Hassler, M. Klatt, K. Kress, S. Lachnit, M. Pafla, F. Schlor, M. Sharinghousen, C. Spener, and M. Steinmaier.

References

- Baburov, Y. (2010): python-readability. <https://github.com/buriy/python-readability>
- Geva, R. (2016): article-date-extractor. <https://github.com/Webhose/article-date-extractor>
- Kohlschütter, C., P. Fankhauser, and W. Nejdl (2010): Boilerplate detection using shallow text features. In: *Proceedings of the third ACM international conference on Web search and data mining* (pp. 441–450). ACM.
- Kouzis-Loukas, D. (2016): *Learning Scrapy*. Packt Publishing Ltd.
- Labs, G. (2016): Goose – Article Extractor. <https://github.com/GravityLabs/goose>
- Lewis, D. D., Y. Yang, T. G. Rose, and F. Li (2004): Rcv1: A new benchmark collection for text categorization research. In: *Journal of machine learning research*, 5 (Apr), 361–397.
- Meschenmoser, P., N. Meuschke, M. Hotz, B. Gipp (2016): Scraping Scientific Web Repositories: Challenges and Solutions for Automated Content Extraction. In: *D-Lib Magazine*, 22 (9/10).
- Ou-Yang, L. (2013): Newspaper: Article scraping & curation. <http://newspaper.readthedocs.io/en/latest/>
- Paliouras, G., A. Mouzakidis, V. Moustakas, C. Skourlas, C. (2008): PNS: A personalized news aggregator on the web. In: *Intelligent interactive systems in knowledge-based environments* (pp. 175–197). Springer.

Identification and Analysis of Media Bias in News Articles

The Impact of Objectivity and Believability on Corporate Decision Making and Performance

*Felix Hamborg¹, Norman Meuschke¹,
Akiko Aizawa², Bela Gipp¹*

¹University of Konstanz, Germany
[first.lastname]@uni-konstanz.de

²National Institute of Informatics, Tokyo, Japan
aizawa@nii.ac.jp

Abstract

Depending on the news source, a reader can be exposed to a different narrative and conflicting perceptions for the same event. Today, news aggregators help users cope with the large volume of news published daily. However, aggregators focus on presenting shared information, but do not expose the different perspectives from articles on same topics. Thus, users of such aggregators suffer from media bias, which is often implemented intentionally to influence public opinion. In this paper, we present NewsBird, an aggregator that presents shared and different information on topics. Currently, NewsBird reveals different perspectives on international news. Our system has led to insights about media bias and news analysis, which we use to propose approaches to be investigated in future research. Our vision is to provide a system that reveals media bias, and thus ultimately allows users to make their own judgement on the potential bias inherent in news.

Keywords: news aggregation; framing; content analysis; media bias

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 224–236.

1 Introduction

The Internet has allowed people to gather knowledge, form their own views, and engage in society (Mossberger, Tolbert & McNeal, 2007). Unrestricted access to unbiased information is crucial for understanding current topics. News articles are the primary source of such information, and thus of particular importance. However, the coverage of media outlets often exhibits a bias, commonly called *media bias*, e.g., due to influencing factors such as the ownership and the source of income of the media outlet, as well as political interference, lobbying, or ideological focus (University of Michigan, 2014). Not only developing countries, but also developed nations are subject to media bias, e.g., in the USA six corporations control 90% of the media (Insider, 2014), which results in a high chance of media manipulation (Esser, Reinemann & Fan, 2001). Trust in media is at a historical low. Less than half of the readers in USA trust media and think it is objective (GALLUP, 2015). Table 1 shows different headlines of two news articles reporting on the same event from 7 November 2014, during the Ukraine crisis.

Table 1: Different headlines about the same event

Publisher	Headline
(CNBC, 2014)	Tank column crosses from Russia into Ukraine: Kiev military
(RT, 2014)	Moscow to Kiev: Stick to Minsk ceasefire, stop making false invasion claims

While Western media, such as CNBC, reported that Russian tanks crossed the Ukrainian border, Russian media, such as RT, primarily portrayed these reports as false claims or did not report about the event. The content and tone of the two articles differs just as strongly as the two headlines suggest. We assume that readers' perceptions of the actual event will differ significantly depending on which article they read.

Also, consider the following headlines of the *same article* published in the New York Times (NYT). The NYT author has covertly altered the article, including its headline, from mainly complimentary (first row) to criticizing back then U.S. president candidate Sanders (second row, which is online now).

Table 2: Different headlines of the same article

Publisher	Headline
(Steinhauer, 2016a)	Bernie Sanders Scored Victories for Years Via Legislative Side Doors
(Steinhauer, 2016b)	Via Legislative Side Doors, Bernie Sanders Won Modest Victories

The author also rewrote several sentences, which initially had a positive tone, to subsequently reflect a dismissive tone. For instance, they removed a complimentary quote on his “very successful strategy”, and replaced it with a paragraph criticizing Bernie Sanders.¹

Reading articles that portray different perspectives on an event helps broaden the narrow and distorted perspective induced by media bias. News aggregators, such as Google News, enable readers to quickly get an overview of the large news landscape. However, most readers only consult a small subset of the available set of sources (Newman, Levy & Nielsen, 2015). Reasons include the overwhelming number of sources, language barriers, or out of habit. Established systems provide no support for showing the different perspectives within related articles. Recent efforts aim to fill this gap and reduce the effects of media bias, but suffer from practical limitations, such as being fine-tuned to one news category, or relying heavily on user input (Munson, Zhou & Resnick, 2009; Park et al., 2009).

After giving an overview of media bias, news aggregation, and news analysis, we introduce research questions that are motivated by weaknesses of state-of-the-art systems to support users’ awareness of media bias and its mitigation. Afterwards, we discuss approaches that address these questions. Finally, we present *NewsBird*, an aggregator that focuses on identification and presentation of shared and different information to reduce media bias effects.

¹ The early version was preserved by the Wayback Machine (Murphy, Hashim & O’Connor, 2007) available at <https://archive.org/web/>.

2 Related work

One of the main effects of media bias is the change of people's awareness and perception of topics (Siemens, 2014), which becomes critical for public issues, such as elections (Bernhardt, Krasa & Polborn, 2008). Reasons for biased news coverage range from intentional bias, e.g., governments and companies influencing publishers in their favor (Besley & Prat, 2002), to rather unintentional bias, e.g., due to the influence of *news values* (Harcup & O'Neill, 2001). One popular example is the NYT's preference for U.S. president candidate Hillary Clinton, which may have resulted for instance in the changes illustrated in table 2. Due to the journalistic requirement of objectivity, media bias is typically rather inconspicuous (Park et al., 2009), e.g., bias is embodied through story selection or placement (front-page story vs. on another page) or word choice ("coalition forces" vs. "invasion forces").

The workflow of news aggregators typically includes article crawling and extraction, summarization of related articles, and visual presentation (Evans, Klavans & McKeown, 2004). However, none of the established systems focuses on revealing differences among related articles (Park et al., 2009). Thus, their users may remain unaware of media bias and possibly different perspectives (Bui, 2010; Wanta, Golan & Lee, 2004).

Recent research efforts have attempted to fill this gap, and aim to reduce the effects of media bias. For instance, *NewsCube* divides a group of related articles using extracted keywords and hierarchical clustering. The resulting subgroups represent different semantic propositions of the main topic (Park et al., 2009). Another approach makes use of users' feedback, who set articles in relation to one another (Park et al., 2011). *Sidelines* counts links from pre-categorized blogs, e.g., liberal vs. conservative, pointing to an article to determine its political orientation. Afterwards, it selects politically opposing articles (Munson et al., 2009).

While these and other systems can reduce the effects of media bias by broadening readers' understanding of news topics, they suffer from several practical limitations. First, they are restricted to the analysis of one news category, i.e., politics (ibid.; Park et al., 2009). Additionally, some of them rely on manually built knowledge bases (Munson et al., 2009; Park et al., 2011).

A major reason for the difficulty of aggregators to show the effects of media bias is that current NLP methods perform especially poorly in identi-

fyng semantic differences in news (Park et al., 2009). In summary, classic NLP techniques that commonly rely on statistics “[...] are just a first step towards natural language understanding” (Cambria & White, 2014). This makes the analysis of news challenging, since semantic differences are often encoded subtly due to journalistic objectivity (Gauthier, 1993).

3 Research questions

This section derives the motivation and goals of our research from the findings of our literature review, which was briefly summarized in section 2. Afterwards, we discuss approaches to address our goals. While the proposals from this section are goals and future work, section 4 describes the current system.

3.1 Research motivation

The fundamental question that motivates our research is how a news analysis system can help users to become aware of media bias and thus reduce the effects of unidentified media bias. While news aggregators do not focus on exposing media bias, aggregators are still a first step in this direction given that they allow users to quickly get an overview of the news landscape. We envision a system that reuses news aggregation design and methods, but is additionally capable of the following:

- (T1) Reveal which news outlets portray which perspectives on what topics.
- (T2) Abstract and aggregate higher level knowledge from (T1), e.g., present possible factors that influence coverage. For instance, different interests in different countries or who is the owner of an outlet.
- (T3) To identify copy-editing and undocumented article changes: enable temporal and contextual analysis of topics to reveal which publishers are influenced by others and how.
- (T4) Provide a platform for users to collectively view and read different perspectives on topics, set publishers in relation, and judge their objectivity, currentness, etc. Such relations could express publishing (dis-)similarity or copy-editing flows.

3.2 News overview

Figure 1 depicts a visualization for a single topic that addresses (T1) from section 3. The upper section briefly summarizes the news topic similar to established news aggregators. The lower section of the view shows the most contrasting perspectives from articles on the selected topic. For geo-based news, such as international topics, a world map (right) visualizes from where these views originate. For other news categories, we will research further supporting views. Information related to (T2) will also be displayed in this view to help users identify the context of this topic and the presented perspectives, e.g., the news outlet that portrays a critical perspective on a green energy topic is owned by a company affiliated with the oil industry.

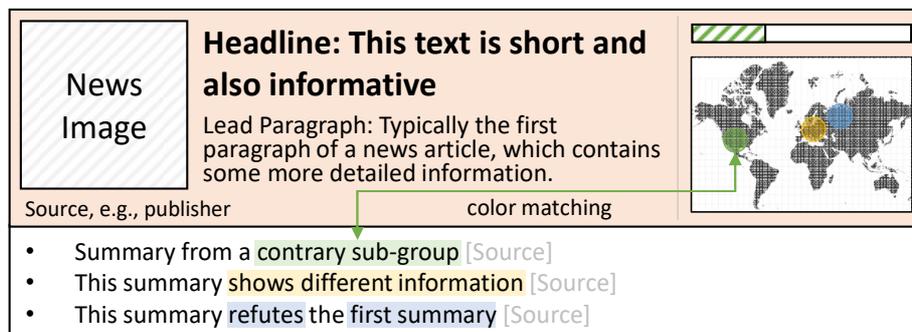


Fig. 1 News topic view for broadened news understanding

3.3 Analysis of news development and relations

We propose a second visualization depicted in figure 2 to analyze temporal and contextual relations between articles, such as if one article is mainly copy-edited from other previously published articles, and if yes, from which article, and which sections were copied or changed. The visualization also shows how an article's content and tone changes over time (cf. semi-transparent popup in fig. 2). This way, our system would help to identify cases of undocumented changes, as shown in table 2. To prove such changes, we plan to use decentralized trusted timestamping (Gipp, Meuschke & Gernandt, 2015).

The visualization in figure 2 addresses (T3). Such dependencies are not visible in current aggregators. The visualization would enable users to explore the origins of an article, i.e., from which an article may have been derived. Each article is represented with a colored box. Dependent articles, e.g., copy-edited from another, are connected with a line and colored in a similar shade.

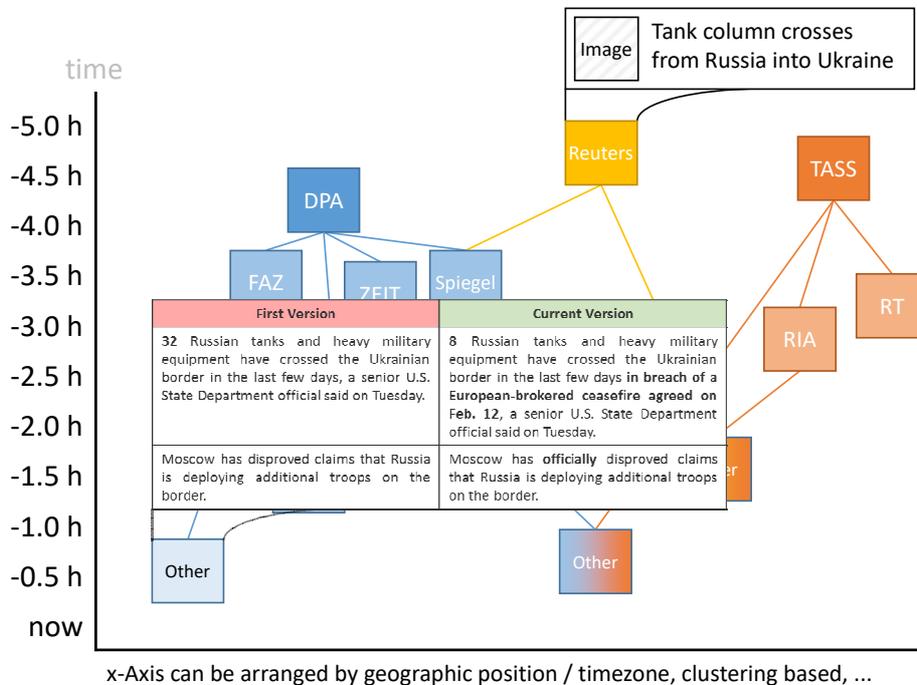


Fig. 2 Visualization for temporal and contextual analysis

3.4 General design of the news analysis

To acquire the information necessary for previously described visualizations, we propose a design that enables the system to analyze any news category, such as international news, finance, etc. While following the typical aggregator workflow (cf. sect. 2), we insert an additional task after topical grouping: *group refinement*. The goal of this task is to refine a group of related articles into subgroups that have a maximized expected diversity, i.e., represent different perspectives on the topic. Our hypothesis is that for each news category there is at least one *dimension* that achieves high diversity among resulting subgroups created by that dimension, e.g., in international news or conflicts often the involved countries have different perspectives on the topic as was demonstrated in table 1. We want to investigate for other categories which dimensions are effective in splitting articles into diverse subgroups.

Figure 3 depicts the system's group refinement workflow, which takes a group of related articles and finds subgroups that have a maximized expected diversity, estimated by the *d-score*. First, our system generates subgroup

candidates for each dimension. The system estimates the subgroups' diversity using the d-score, for which it uses user-provided relations and a method we will term *document similarity analysis* (DSA). DSA measures the semantic (dis-)similarity between two documents by comparing features such as extracted keywords, syntax, hyperlinks, images, and other elements common to news. We will gather user-provided relations directly in our system by facilitating such feedback (Park et al., 2011). Users can provide publisher and article relations, and quantified information on news, such as rate their objectivity and currentness (T4) (cf. *ibid.*). Finally, our system selects subgroups that yield the highest d-score. The results are then presented in one of the presented visualizations.

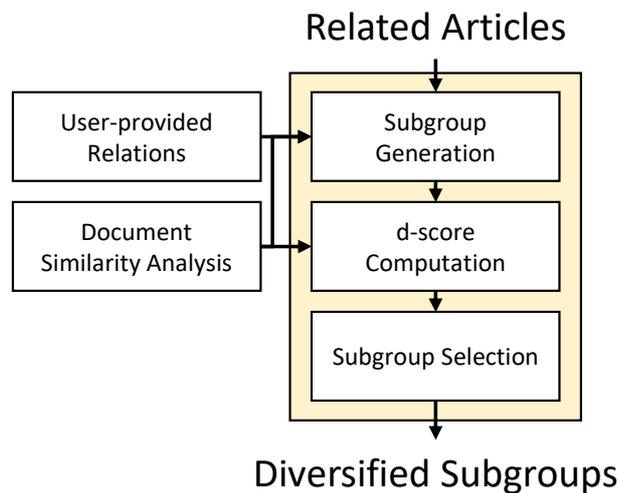


Fig. 3 Group refinement workflow

4 *NewsBird*: Different perspectives in international news

We developed *NewsBird* as a news aggregator that additionally addresses the goals described in section 3.1. Our aggregator shows both the shared and the opposing perspectives expressed in topically related news articles. *NewsBird* currently focuses only on international news, which we consider a prime example of media bias as displayed in table 1 and figure 4. We plan to general-

ize the design to support further news categories, and add visualizations that help users investigate the effects of media bias and get broad news understanding (cf. sect. 3). Therefore, we will use the findings we obtained during the development of the prototype and from an expert case study.

Figure 4 depicts NewsBird’s design, which is built around the idea of analyzing and visualizing news topics in a matrix to reveal their differences. In a onetime or regular process, the first two stages *data gathering* and *articles extraction* are performed to create or update the database by insertion of news articles. We used a non-disclosed dataset that contains 1.6 million articles gathered from almost 4,000 publishers from over 100 countries in October and November 2014. The dataset stems from the European Media Monitor (Atkinson & Goot, 2009). In future, we plan to use news-please to be able to aggregate and analyze articles reporting on further events (Hamborg et al., 2017).

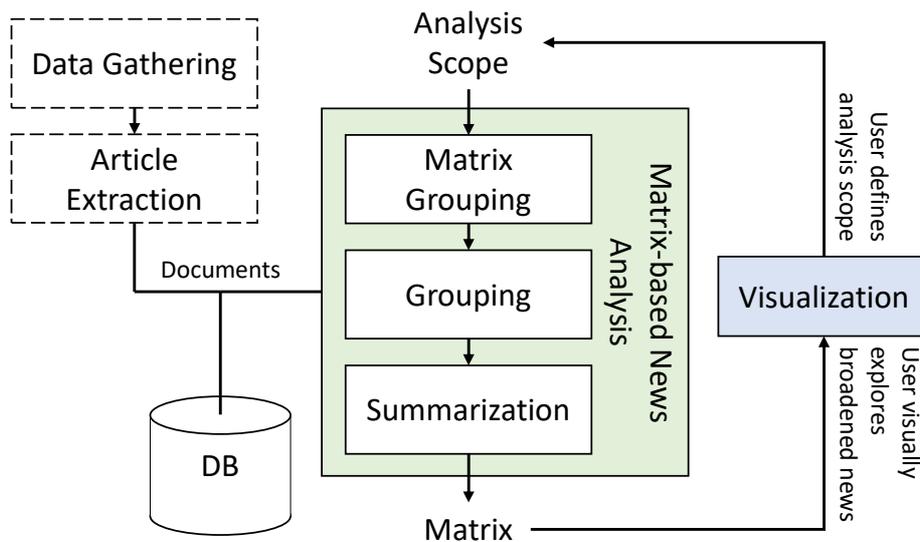


Fig. 4 Analysis design of NewsBird

First, in the *visualization* users define the *analysis scope*, which contains query date and two matrix dimensions including their specific values. To support the analysis of international news, NewsBird’s dimensions are the *publisher* and *mentioned country*. For instance, the cell of publisher country Russia and mentioned country Ukraine, hereafter denoted with RU-UA, contains all articles that have been published in Russia and mention Ukraine. This way, NewsBird reveals what is stated in which country about another

country, e.g., RU-UA contains articles a Russian news reader might read about the Ukraine.

The first analysis task is *matrix grouping*, which spans a matrix over both dimensions, creates corresponding cells, and queries articles for each cell. Afterwards, the *grouping* task finds groups of related articles, i.e., which report on the same topic. NewsBird finds these topics and subsequently calculates which articles are related by performing LDA on the articles across all matrix cells. The *summarization* generates summaries for 1) each topic and each cell 2) on all the cell's topics, and 3) for each of the topics of the cell. NewsBird uses the tokens' TF-IDF weights and cumulates them on sentence level.

Finally, NewsBird's *visualization* presents these analysis results to the user. Similar to established news aggregators, NewsBird allows users to get a news overview by showing them a list of topics and related articles. The key difference is our matrix visualization, which enables a comparative news exploration by showing both shared and opposing information in parallel as depicted in figure 5.

		Mentioned Countries			
		UA	RU	GB	DE
Publisher Countries	RU	Foreign Policy Adviser Says Russia Committed to Peace Process in East Ukraine	Ukraine Crisis, Sanctions Against Russia Not on G20 Agenda in Australia: Russian Sherpa	Cameron Says Britain Will Pay Only Half of \$2.6 Bln EU Surcharge	Berlin wall: the symbol of Cold War as an art object
	GB	Ukraine crisis: Kiev accuses Russia of military invasion after 'tanks cross border'	Tank column crosses from Russia into Ukraine – Kiev military	Cameron has warned there will be a „major problem“ if Brussels insists on Britain paying its \$2.6 bn	Fall of the Berlin Wall: „Our tears of frustration turned to those of joy“
	DE	Kyiv calls Berlin amid Russian incursion reports	Kyiv: 32 tanks enter Ukraine from Russia	Britain allowed to halve EU budget bill	Germany's east still lags behind
	US	Ukraine accuses Russia of sending in dozens of tanks	Ukraine accuses Russia of sending in dozens of tanks	Britain finds deal with EU over controversial bill	AP WAS THERE: The Berlin Wall crumbles

Fig. 5 Comparative news exploration in NewsBird

Figure 5 illustrates the efficacy of NewsBird's visualization on the exemplary topic of the Ukraine crisis from 7 November 2014, which was introduced in table 1. Note how the matrix reveals that media in the analyzed Western countries (rows) primarily reports that Russian tanks have invaded

Ukraine (in both Russia and Ukraine columns). However, the cell RU-UA shows that Russian media's most frequent topic involving the Ukraine is Russia's commitment to the peace process in the Ukraine.

We conducted an expert case study with one participant where we collected observational and think aloud data as well as information from eye gaze tracking and a follow-up interview to assess the subjective experience. Given the early stage of this project, we were mainly interested in whether users find NewsBird helpful to become aware of media bias and understand resulting differences in news coverage. We used three real-world news scenarios, including the Russian tank invasion on 7 November 2014, which we briefly introduced in section 1. We found that NewsBird enables users to get an overview of the current news situation similarly to established news aggregators. Furthermore, the user reported that using NewsBird he became aware of media bias and how this can lead to differences in news coverage for the first time.

5 Conclusion

In this paper, we introduced NewsBird, a news aggregator that aims to address practical limitations of established systems. NewsBird uses a two-dimensional matrix spanned by publisher and mentioned countries. This way, related articles are split into subgroups that have a large expected diversity and reveal differences in news coverage. While the current prototype is limited to the category of international news, our expert case study has shown that NewsBird already enables users to identify the effects of media bias and thus broaden users' news understanding. Furthermore, NewsBird has helped us identify directions for further research. The main goal will be the generalization of the current workflow to further news categories. We proposed multiple concepts for investigation, such as visualizations that reveal article and publisher dependencies, and a method to show temporal and contextual changes in news articles. To implement these aims, we will investigate and enhance NLP methods, e.g., to extract entities from news texts and measure article similarity by using both text and non-text features of news articles. This will help users in identifying and comparing media bias in current news topics, and thus ultimately reduce the effects of media bias in news coverage.

References

- Atkinson, M., and E. Van der Goot (2009): Near real time information mining in multilingual news. In: *Proceedings of the 18th international conference on World Wide Web* (pp. 1153–1154). ACM.
- Bernhardt, D., S. Krasa, and M. Polborn (2008): Political polarization and the electoral effects of media bias. In: *Journal of Public Economics*, 92 (5), 1092–1104.
- Besley, T. J., and A. Prat (2002): Handcuffs for the grabbing hand? Media capture and government accountability. CEPR Discussion Paper.
- Bui, C. (2010): How online gatekeepers guard our view: News portals' inclusion and ranking of media and events. In: *Global Media J.*, 9 (16), 1–41.
- Cambria, E., and B. White (2014): Jumping NLP curves: a review of natural language processing research. In: *Computational Intelligence Magazine, IEEE*, 9 (2), 48–57.
- CNBC (2014): Tank column crosses from Russia into Ukraine: Kiev military. <http://www.cnbc.com/id/102155038>
- Esser, F., C. Reinemann, and D. Fan (2001): Spin Doctors in the United States, Great Britain, and Germany Metacommunication about Media Manipulation. In: *The Harvard International Journal of Press/Politics*, 6 (1), 16–45.
- Evans, D. K., J. L. Klavans, and K. R. McKeown (2004): Columbia newsblaster: multilingual news summarization on the Web. In: *Demonstration Papers at HLT-NAACL 2004* (pp. 1–4). ACL.
- GALLUP (2015): Americans' Trust in Media Remains at Historical Low. <http://www.gallup.com/poll/185927/americans-trust-media-remains-historical-low.aspx>
- Gauthier, G. (1993): In defence of a supposedly outdated notion: The range of application of journalistic objectivity. In: *Canadian Journal of communication*, 18 (4), 497–505.
- Gipp, B., Meuschke, N., & Gernandt, A. (2015): Decentralized Trusted Timestamping using the Crypto Currency Bitcoin. In: *Proceedings of the iConference 2015* (to appear). Newport Beach, CA, USA. <http://ischools.org/the-iconeference/>
- Hamborg, F., Meuschke, N., Breitingner, C., & Gipp, B. (2017): news-please: A Generic News Crawler and Extractor. In: M. Gäde, V. Trkulja and V. Petras (eds.): *Everything Changes, Everything Stays the Same? Understanding Information Spaces*. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Hülsbusch, pp. 218–223.
- Harcup, T., & O'Neill, D. (2001): What is news? Galtung and Ruge revisited. In: *Journalism studies*, 2 (2), 261–280.

- Insider, B. (2014): These 6 Corporations Control 90% of the Media in America. <http://www.businessinsider.com/these-6-corporations-control-90-of-the-media-in-america-2012-6>
- Mossberger, K., Tolbert, C. J., & McNeal, R. S. (2007): *Digital citizenship: The Internet, society, and participation*. MIT Press.
- Munson, S. A., Zhou, D. X., & Resnick, P. (2009): Sidelines: An Algorithm for Increasing Diversity in News and Opinion Aggregators. ICWSM.
- Newman, N., Levy, D. A., & Nielsen, R. K. (2015): Reuters Institute Digital News Report 2015. Available at SSRN 2619576.
- Park, S., Kang, S., Chung, S., & Song, J. (2009): NewsCube: delivering multiple aspects of news to mitigate media bias. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 443–452). ACM.
- Park, S., Ko, M., Kim, J., Choi, H., & Song, J. (2011): NewsCube 2.0: An Exploratory Design of a Social News Website for Media Bias Mitigation. In: *Workshop on Social Recommender Systems*.
- RT (2014): Moscow to Kiev: Stick to Minsk ceasefire, stop making false ‘invasion’ claims. <http://rt.com/news/203203-ukraine-russia-troops-border/>
- Siemens, G. (2014): Connectivism: A learning theory for the digital age. http://www.itdl.org/journal/jan_05/article01.htm
- Steinhauer, J. (2016a): Bernie Sanders Scored Victories for Years via Legislative Side Doors. <https://web.archive.org/web/20160314164825/http://www.nytimes.com/2016/03/15/us/politics/bernie-sanders-amendments.html?partner=rss&emc=rss>
- Steinhauer, J. (2016b): Via Legislative Side Doors, Bernie Sanders Won Modest Victories. <https://web.archive.org/web/20160314164825/http://www.nytimes.com/2016/03/15/us/politics/bernie-sanders-amendments.html?partner=rss&emc=rss>
- University of Michigan (2014): News Bias Explored – The art of reading the news. <http://umich.edu/~newsbias/index.html>
- Wanta, W., Golan, G., & Lee, C. (2004): Agenda setting and international news: Media influence on public perceptions of foreign nations. In: *Journalism & Mass Communication Quarterly*, 81 (2), 364–377.

Using Sessions from Clickstream Data Analysis to Uncover Different Types of Twitter Behaviour

Florian Meier, Johannes Aigner, David Elweiler

Information Science Group
University of Regensburg, Germany
{florian.meier, johannes.aigner, david.elweiler}@ur.de

Abstract

While much is known about how Twitter is used for specific tasks or by particular groups of users, we understand surprisingly little about how the service is used generally on a daily basis. To learn more about general Twitter behaviour we perform a cluster analysis on a rich set of longitudinal interaction log data describing interactions 44 users had with the Twitter website over a 5 month period. We report on and interpret 5 clusters representing common usage patterns with the service

Keywords: Twitter; clickstream data; clustering; information behaviour

1 Introduction

No Social Media application has attracted more research attention than the micro-blogging service Twitter. Many research areas make use of Twitter data for diverse purposes. These include linguistic analysis (Owoputi et al., 2013), sentiment analysis (Nakov et al., 2016), event detection (Sakaki, Okazaki & Matsuo, 2010) or to understand behaviour with Twitter during specific events, such as science conferences (Wen, Trattner & Parra, 2014) or

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 237–250.

presidential debates (Lin et al., 2013). Few studies, however, have investigated how Twitter is used on a day-to-day basis. We do not know, for example, what a typical Twitter session looks like, which features or sub-pages of the site are important or how Twitter behaviour can be characterized in general. This belies the fact that “[u]nderstanding how users behave when they connect to social networking sites creates opportunities for better interface design, richer studies of social interactions, and improved design of content distribution systems“ (Benevenuto et al., 2009: 49). Our main focus in this work is to understand what Twitter session behaviour can tell us about different types of Twitter use. To this end we leverage rich behavioural clickstream data from a naturalistic setting to gain insight into Twitter users’ behaviour at a session level. 44 users provide unrestricted access to their interaction data with the Twitter website over a period of 5 months. We first identify user session then, focusing on 57 specific interactions (features), we perform a cluster analysis to identify common session types.

Our specific contributions are: (i) We compare 560 clustering approaches using different clustering algorithms, cluster sizes and input parameters to find the best combination for the data, (ii) We identify five primary clusters of Twitter sessions revealing different types of Twitter behaviour, (iii) We characterize these cluster types in detail by referring to the session features they exhibit, (iv) Finally, we investigate whether the time of day has an influence on the way Twitter users behave.

2 Related work

The two lines of research most relevant to the work described in this paper are: (i) studies investigating general Twitter behaviour and (ii) studies analysing social media clickstream data. We give a short review of both.

(i) Java et al. (2007) were amongst the first to study Twitter behaviour. They created a taxonomy of intentions when using the service, as well as types of users. Subsequent work has investigated the usage of certain Twitter features including @-Mentions (Honeycutt & Herring, 2009; Wang, B. et al. 2013), hashtags (Lin et al., 2013), favouriting/liking (Meier, Elswailer & Wilson, 2014), forwarding of messages via retweets (RT) (Boyd, Golder & Lotan, 2010) and how the Twitter search is used (Teevan, Ramage & Morris,

2011). Other work has investigated how certain groups of users such as celebrities or academics make use of Twitter (Marwick & Boyd, 2011; Wen, Trattner & Parra, 2014). To our knowledge no studies exist, which have investigated Twitter behaviour generally, i.e. to determine what a typical Twitter session looks like or how behaviour can vary.

(ii) Click-through or click-stream data refers to data that records and stores user interaction with an application in a highly detailed manner (Dumais et al., 2014). In the context of social media applications Benevenuto et al. used the HTTP request response pairs from a Brazilian social network aggregator to compare user behaviour characteristics from several thousand users of four different social media applications (*Orkut*, *MySpace*, *Hi5*, *LinkedIn*) (Benevenuto et al., 2009). They discovered that up to 92% of user interaction is *latent* i.e. browsing behaviour (ibid.: 58). This reveals that crawled data or data collected via APIs accounts for only a tiny fraction of user behaviour. Schneider and colleagues published similar work studying four different social media applications (*Facebook*, *LinkedIn*, *Hi5*, *StudiVZ*) and finding similar outcomes (Schneider et al., 2009). Wang et al. use click-stream data¹ from the Chinese social media application *Renren* to build click models capable of distinguishing between normal users and malicious users or bot accounts (Wang, G. et al., 2013). Similar to our work Wang et al. are using unsupervised clustering approaches for user behaviour analysis (Wang, G. et al., 2016). However, they focus on designing, implementing and testing a framework for visualizing behavioural clusters and not describing the clusters in detail. Although these studies provide initial insights into social media use, most of the analysis is high-level and lacks a detailed description of usage patterns.

Finally, the work by Buscher et al. is relevant from a methodological perspective (Buscher et al., 2013). In their work they also used clustering on clickstream data for SERP page interaction to identify different behavioural clusters. Their approach and discussion of results inspired our own analyses.

¹ 7 million clicks from 16,000 users in a period of 2 months

3 Study methodology

3.1 Collecting clickstream data

The Twitter API offers simple access to behavioural data for the service. Data sourced in this way is limited in that it is biased towards active engagement and omits latent interactions, such as tweet consumption and navigation of sub-pages. To get a detailed picture of how users interact with Twitter requires a richer and more detailed set of interactions to be captured. This was the motivation for designing and implementing an extension for the Chrome web browser. The extension records user interaction in a highly detailed manner. Besides metadata from viewed or interacted tweets the extension recorded the timestamp, the type of interaction, the context (URL) in which it occurred (e.g. the users own timeline), as well as detailed information on mouse movement (e.g. clicking or scrolling).

3.2. Participants

From end of March 2015 to the end of August 2015 44 users downloaded and installed the extension from the Chrome Web Store. The link to the extension was passed on via Twitter posts, e-mail lists and internet fora. Our aim was to recruit a sample of users that is as heterogeneous as possible with respect to their account statistics in order to investigate Twitter session behaviour from a diverse range of different users. Table 1 shows the account information from 27 of the users, who could reliably be mapped to an account and consented to being identified. All account statistics are highly skewed and long tailed, which hints to varying Twitter behaviour among the participants. During the study about 180,000 events were recorded and participants viewed 270,000 tweets.

Table 1: Twitter Account statistics for the study participants

(<i>n</i> = 27)	min	max	mean	median
Tweets (<i>n</i>)	0	73,780	5,637	59
Following Count (<i>n</i>)	4	1,041	203	57
Follower Count (<i>n</i>)	0	1,139	202.5	39
Favourites Count	0	44,520	1,796	18.50
Registered Years (years) (NA = 12)	2	6	4.6	5

3.3 Event categorization and session segmentation

Several data preparation steps were performed before analysing the data. First, low-level single events were categorized into higher-level event categories based on the context (URL) in which they occurred. The first event in figure 1 shows a click on a tweet occurring on the own timeline (URL: <http://twitter.com>), which was categorized as a TIMELINE² event. Certain types of events have fixed event categories regardless in which context they occurred. E.g. a search query is always classified as an event from event category SEARCH.

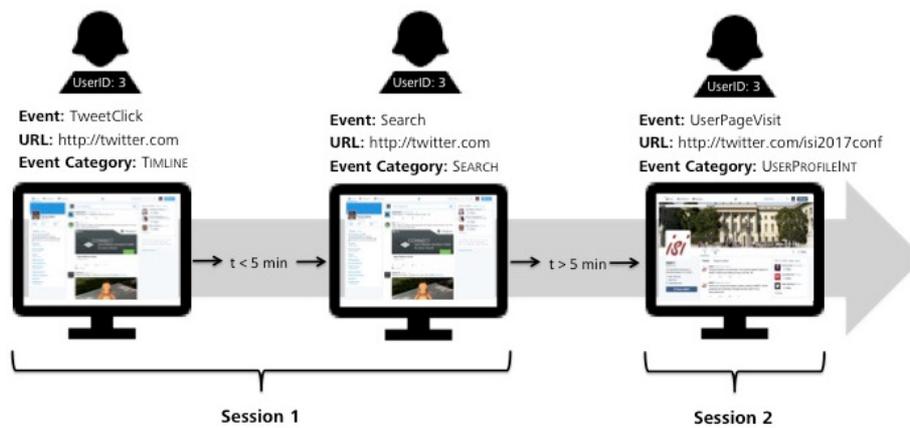


Fig. 1 Visualization of important data preparation steps

In a next step, the stream of events was split into sessions by using five minutes of user inactivity as a threshold for the start of a new session (fig. 1) an established approach in the literature (Schneider et al., 2009).

4 Session characteristics and clustering experiment

Using five minutes inactivity as a session threshold results in an overall session count of 4573 sessions for all users. On average a session contains 38.71

² From this point further event categories are set in small caps.

events (min = 2, max = 2155, median = 11, IQR = 31) and lasts slightly over 4 minutes (median = 83 sec, IQR = 270 sec). However, sessions can be as short as only 1 sec (min) or even last for 3 hours and longer (max = 11,080 sec). About 25% of participants used Twitter for 2 sessions per day and 40% of the participants had one session per day during the study period. However, the maximum lies at 10 sessions per day. These stark differences are a first hint that users behave differently and have different motivations and goals in mind when using Twitter.

We investigate whether session features described in figure 2 characterize sessions in a way that they can be used to discriminate sessions from each other and reveal different types of Twitter usage patterns. To this end we applied an unsupervised learning approach to the sessions in order to allow for different patterns of behaviour to emerge from the data. We empirically assessed the best clustering approach for our data by comparing different algorithms, cluster sizes ($k = 2$ to 15) and input parameters. We compared the following five clustering algorithms: (i) K-Means (ii) Fuzzy C-Means (iii) Partitioning around Medoids (PAM) (iv) Hierarchical Clustering (Ward D).

Finally, we added (v) Spherical K-Means with repeated bisection method as this algorithm worked well on similar data as shown by Buscher and colleagues (Buscher et al., 2013). As input parameters we selected the following five features or rather group of features from all possible features describing a session (cf. fig. 2). (i) EventsTotal (ii) Duration (iii) SessionBusynessIndex (Events per Duration) (iv) Events Per Event Category (Count) (v) Events Per Event Category (%).³ We combined at least two of the features and used them as input parameters for the clustering process in a stepwise manner. This process resulted in a total of 560⁴ clustering results in the form of average silhouette values (avgSIL), a measure of cluster validity introduced by Rousseeuw (1987). The silhouette value is a measure combining cohesion (similarity or relatedness of objects within a cluster) and separation (distinctness or separation of a cluster from other clusters) into a single measure of cluster quality (Kumar, 2005: 536). It is a useful measure as not only every single data point has a SIL value, but the SIL can also be used to measure the soundness of single clusters within a clustering result as well as the cluster result in total by averaging the SIL values for every cluster, as we did in our experiment (ibid.: 542). Using the Spherical K-Means algorithm and using

³ See features 4 to 15 from Figure 2.

⁴ 5 algorithms * 14 different cluster sizes * 8 input parameter combinations = 560

the SBI and the relative number of events per event category as input parameters achieved the most promising result. The highest average SIL (0.74) can be detected at a cluster size of $k = 5$.

<i>Session Features</i>	
EventsTotal	Number of events per session
Duration	Session duration in seconds
BusynessIndex	Number of events per duration
TimelineEvents†	Number of events per context Timeline
NotificationsEvents†	Number of events per context Notifications
UserProfileEvent†	Number of events on user profiles
OwnProfileEvent†	Number of events on own profile
FFMRelationsEvents†	Number of events for Following-Follower-Mention
SearchEvents†	Number of events per context search/number of search events
FavouritesEvents†	Number of events per context favourites list
SingleTweetEvents†	Number of events per single Tweet view
TweetingEvents†	Number of Tweets posted per session
DirectmessageEvents†	Number of direct message interactions
ListEvents†	Number of events on lists
OtherEvents†	Number of events on other sub-pages (e.g. Twitter API site)
TweetsSeen	Number of Tweets seen by user in session
MUR	Message uncertainty ratio (Tweets clicked/distinct Tweets clicked)
RTCountSession*	Average and maximum number of RTs observed over all Tweets seen
FAVCountSession*	Average and maximum number of FAVs observed over all Tweets seen
TweetPosition*	Average and maximum position observed over all Tweets seen
TweetAge*	Average and maximum age of Tweet observed over all Tweets seen
TweetVistime*	Average and maximum Tweet visibility till click
TweetsBuffer†	Number of Tweets added to stream while scrolling
TweetsClick†	Number of Tweets clicked
TweetsHover†	Number of Tweets hovered
ScrollSummary	Number of continuous scrolling activities
DownScrollTotal	Down scroll distance in px
UpScrollTotal	Up scroll distance in px
ScrollIndex	Total pixels scrolled/scroll summary events
MouseMoveCount	Number of distinct mouse movements
MouseDistance*	Average and maximum distance covered by cursor in px
AvgMouseActivityIndex	Site area (square px) covered by number of mouse movements
ProfileVisits	Number of user profile sites visited
TotalProfileTime	Number of time spent on user profiles
DistinctProfilesVisited	Number of distinct user profiles visited
PUR	Person uncertainty ratio

Fig. 2 Features as result of the feature engineering process that characterize Twitter sessions. Features tagged with † have absolute and relative values. Features tagged with * have average and max values. Figure taken from Meier & Elswailer (2016: 361).

Additionally, besides looking at the SIL width, one also has to manually check the single SIL values of each cluster of the clustering result to determine the soundness of cluster sizes. Besides one dominant cluster the number of sessions per cluster is balanced and each cluster at least has a SIL value of 0.36. We take these clusters for further interpretation in the following section.

5 Cluster interpretation

The clusters were interpreted by taking a multistep approach. In a first step we investigated which event categories i.e. Twitter subpages were particularly popular compared to the rest of the event categories in the sessions in each cluster. This can be done by interpreting the distribution of events per event category visualized by the boxplots in figure 3. In a next step we derived potential use cases for the dominant event categories in every cluster. We try to think of plausible sequences of events that explain the dominance of these event categories and identify behavioural patterns in the five clusters.

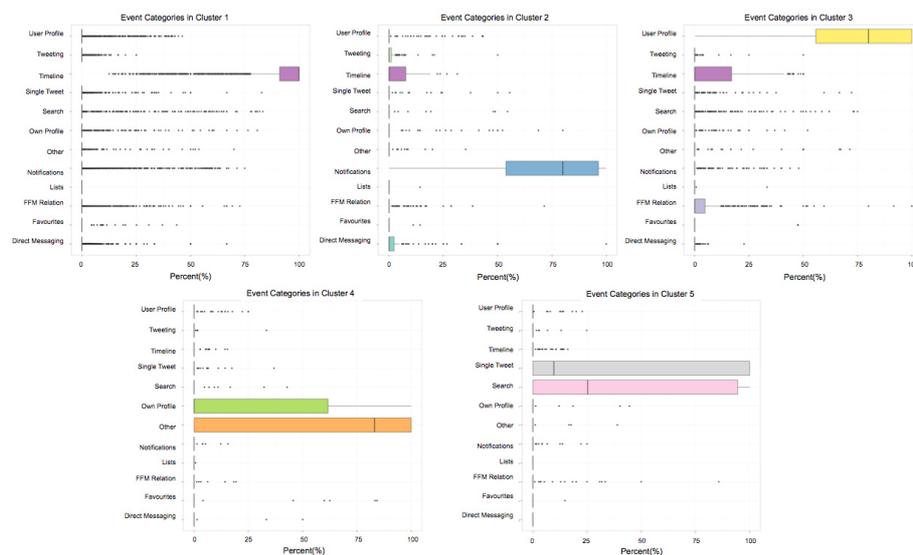


Fig. 3 Proportion of events per event category for sessions in each cluster

Cluster 1, **Surf Timeline**⁵ ($n = 3718$): shows a clear dominance of events from event category TIMELINE. The median of the boxplot lies at 100% revealing that half of all sessions from that cluster have timeline interaction exclusively. Other event categories are dominated by outliers to the top hinting to the fact that within sessions of this cluster other Twitter subpages are rarely or never visited. As most sessions belong to this cluster those can be interpreted to be the most typical Twitter sessions.

⁵ From this point further clusters names have the font weight bold.

Cluster 2, **Check Notifications** ($n = 157$): although the NOTIFICATIONS category is the most dominant, events from other event categories, namely TWEETING, TIMELINE and DIRECT MESSAGING are very popular too. This pattern shapes a coherent picture of users looking at their notifications and being active themselves by either tweeting or sending direct messages.

Cluster 3, **User Interaction** ($n = 492$): most events occur on other user's profile pages (USERPROFILEINT). Compared to other clusters a certain proportion of events from event category FFM RELATIONS are observed. Users clicking on @-Mentions and visiting the profile sites of this user can explain these high proportions. Another combination of both categories could be the action of users un/following accounts by using the button from this users' profile page.

Cluster 4, **Other and Own Profile** ($n = 96$): This cluster shows a dominance of event category OTHER combined with events from category OWN-PROFILEINT. Via their own profile users are able to view their posted tweets or finding links to their favourites list or profile settings. Users who want to get more information on Twitter as a company, look up the Twitter blog, or visit the Twitter API pages all generate events within Twitter subpages belonging to event category OTHER.

Cluster 5, **Information Seeking** ($n = 110$): The last cluster shows a dominance of events from category SEARCH and SINGLE TWEET. It is the most information behaviour related cluster as it reveals users searching for information by issuing queries or clicking on hashtags and investigating tweets and responses to those tweets in more detail by viewing them in the single page view.

In a next step we are using the features introduced in figure 2 to characterize each cluster in more detail and if necessary find differences between them. A highly significant Shapiro-Wilk test for all features and all clusters proves that they are all non-normally distributed as such when testing for significance the non-parametric Wilcoxon test was used.

Most sessions belong to the cluster **Surf Timeline**. Looking at the features from table 2 these sessions tend to be rather short (*Duration (sec)*) and not having many events (*Events (n)*). However, the tweets viewed by the users are fresh (*Tweet Age (hours)*) and on top of the stream (*Tweet Position*). All these facts hint at a quick update check, to look at the latest tweets in the timeline. Other contexts or sub-pages are only rarely visited.

In cluster **Notifications** the dominance of the events in the categories NOTIFICATIONS hints to the behavioural picture of users viewing their notifi-

cation (replies, mentions, favourites and retweets from other users) on their notification page. Moreover, the events in event category TWEETING and DIRECT MESSAGING suggest that they post own tweets or read and send direct messages. Tweeting and sending direct messages could be a reaction to viewed notifications, but using this clustering approach alone we cannot confirm this.

Table 2:

Selected Features for cluster characterization denoted as their median

Meta-cluster Name		Surf Timeline	Notifications	User Profile	Own Profile/Other	Search/Single Tweet
Cluster		#1	#2	#3	#4	#5
Number of Sessions		3718	157	492	96	110
Feature (Median)						
Events	Events (n)	10	14	25	8.5	7
Time	Duration (sec)	73	115	157	80	46.50
	Profile Visit Time (sec)	17.50	7	67	18.50	5
	SBI (Events/Time)	0.24	0.17	0.19	0.15	0.23
	Tweets	Tweet Age (hours)	4	10.66	138.36	161.30
	Tweet Position	1.57	1.40	3.28	1.25	3.83
Clicks	Tweet Clicks (n)	2	1	1	1	1
Hover	Tweets Hovered (n)	4	4	7	2	6
Scroll	Down Scroll (px)	2946	466.67	3128	2390	1100.57
	Up Scroll (px)	100	302	292	660	0
	Scroll Index	395.27	239.50	445.17	506.63	310.40
Mouse	Mouse Move Count (n)	81	50	100	50	50
	Distance in Session (px)	140.85	99.41	153.11	164.86	109.57
	Total Distance in Session (px)	12009.72	5479.53	14502.82	10465.48	7349.22
Event Types	Go to: Direct Message	1596	60	31	4	0
	Go to: Notifications	1537	738	238	7	6
	Go to: Timeline	4842	228	597	32	124

Table 2 suggests that with respect to *Duration (sec)* sessions from cluster **User Interaction** are significantly longer than sessions from any other cluster. Four significant Wilcoxon tests evidence this impression. The table paints a similar picture with regard to the number of events (*Events (n)*). The fact that *Median Profile Visit Time (sec)* is longest (67 sec) for sessions in this cluster is a further hint that user profile interaction is essential for sessions in this cluster. Additionally, it is striking that during those sessions many tweets get viewed (*Tweets Hovered (n)*) and those tweets are much older (*Tweet Age (hours) = 138.36*) compared to the median age of tweets from other clusters. Finally, when looking at all features concerning mouse movement and scrolling one can see that these sessions are very intense and much harder work as those values are all very high. Cluster 3 is the second biggest cluster and as such visiting user profile sites happens quite fre-

quently. Reasons for visiting user profiles could be manifold. Whether it is for checking a users latest tweets, making an informed decision on whether one should un/follow this account, or if it is for re-finding previously seen tweets, which other studies have shown (Meier & Elsweiler, 2016). All these motivations can lead to characteristic behaviour on user profiles.

Besides the median age of viewed tweets being oldest in Cluster 4 no other distinctive features can be observed. However, the small numbers of viewed tweets can be explained by most interaction occurring on subpages of category OTHER on which no tweets are present.

A slightly different trend can be observed concerning the **Information Seeking** cluster. Sessions from this cluster have significantly less events than session from clusters 1, 2 and 3 but are not significantly longer or shorter compared to **Surf Timeline** or **Own Profile / Other**. Moreover, features with respect to scrolling and mouse movement are also less prevalent compared to other clusters. However, more tweets are hovered over (*Tweets Hovered* (n) = 6). This indicates that during search sessions there's a certain focus on consumption of information (i.e. reading tweets) rather than interaction. Moreover, the *Event Types* rows from table 3 indicate that in sessions from cluster 5 no steps to other event categories are taken, thus searching for information seems to be a rather self-contained, separate kind of behavioural pattern.

Finally, we analysed whether time of day has an effect on the occurrence of sessions from certain clusters. We hypothesized that sessions from cluster **Surf Timeline** would rather occur during the day whereas sessions that are more time intensive, like sessions from cluster 2, 3, 4 would rather occur after work i.e. during spare time. We divide sessions into two groups on whether they belong to **Surf Timeline** or not. Further, we look at the period between 6:00 AM and 11:00 PM and split the sessions on whether they started before 2:00 PM or after 2:00 PM. A Chi-Square test couldn't find a significant difference between the distributions ($p = 0.297$). In a next step we built six different groups by splitting the time period into morning, noon and evening. Again pairwise Chi-Square tests comparing the distributions of all possible combinations showed no significant differences. We conclude that time of day does not have an effect on type of Twitter behaviour.

6 Summary and future work

This work presented a characterization of Twitter user behaviour by using clustering on session data from a log study of 44 users in a naturalistic setting. In a clustering experiment we evaluated 560 approaches in which the best approach separated 4573 user sessions into 5 clusters. By using session-related features we were able to identify several typical Twitter behaviours. The most notable three are: (i) sessions from cluster **Surf Timeline** represent the typical Twitter session, which is about 73 sec long, where interaction mostly occurs in the context TIMELINE and the latest tweets are viewed. (ii) The **User Interaction** cluster where interaction mostly occurs on the profile sites of other users and sessions tend to be long and intensive and many old tweets are viewed. (iii) The **Information Seeking** Cluster, where people deliberately search for information.

The User Interaction cluster probably reveals the most surprising behaviour and opens up several possible questions. What are motivations for visiting the profile sites of other users? What tasks are users trying to fulfil when visiting profiles? We mention some, but there are probably many more. There is certainly room for improving the user profile page view in general to support users with the variety of tasks they have. Investigating sequences of low-level events and thus explaining those clusters in even more detail are possibilities for future work.

References

- Benevenuto, F., T. Rodrigues, M. Cha and V. Almeida (2009): Characterizing User Behavior in Online Social Networks. In: *Proc. IMC '09* (pp. 49–62). New York, NY: ACM.
- Boyd, D., S. Golder and G. Lotan. (2010): Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. In: *Proc. HICSS '10* (pp. 1–10). Washington, DC: IEEE CS.
- Buscher, G., R. W. White, S. Dumais and J. Huan. (2012): Large-Scale Analysis of Individual and Task Differences in Search Result Page Examination Strategies. In: *Proc. WSDM '12* (pp. 373–382). New York, NY: ACM.

- Dumais, S., R. Jeffries, D. M. Russell, D. Tang and J. Teevan (2014): Understanding User Behavior through Log Data and Analysis. In: *Ways of knowing in HCI* (pp. 349–372). Springer.
- Lin, Y., D. Margolin, B. Keegan, A. Baronchelli and D. Lazer (2013): #Bigbirds Never Die: Understanding Social Dynamics of Emergent Hashtags. In: *ICWSM '13* (pp. 370–379). Palo Alto: AAAI.
- Marwick, A. and D. Boyd. (2011): To See and Be Seen: Celebrity Practice on Twitter Convergence. In: *The International Journal of Research into New Media Technologies*, May, 17, 139–158.
- Meier, F., D. Elswailer and M. L. Wilson (2014): More than Liking and Bookmarking? Towards Understanding Twitter Favouriting Behaviour. In: *Proc. ICWSM '14*. Palo Alto: AAAI.
- Meier, F. and D. Elswailer (2016): Going Back in Time: An Investigation of Social Media Re-finding. In: *Proc. SIGIR '16* (pp. 355–364). New York, NY: ACM.
- Nakov, P., A. Ritter, S. Rosenthal, V. Stoyanov and F. Sebastiani (2016): SemEval-2016 task 4: Sentiment analysis in Twitter. In: *Proc. SemEval '16*, San Diego, CA, USA.
- Owoputi, O., B. Connor, C. Dyer, K. Gimpel, N. Schneider and N. A. Smith (2013): Improved Part-of-Speech Tagging for Online Conversational Text with Word Clusters. In: *Proc. naacl-hlt '13* (pp. 380–390).
- Rousseeuw, P. J. (1987): Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. In: *Journal of Computational and Applied Mathematics*, 20 (C), 53–65.
- Sakaki, T., M. Okazaki and Y. Matsuo (2010): Earthquake Shakes Twitter Users: Real-Time Event Detection by Social Sensors. In: *Proc. WWW '10* (pp. 851–860). New York, NY: ACM.
- Schneider, F., A. Feldmann, B. Krishnamurthy and W. Willinger (2009): Understanding Online Social Network Usage from a Network Perspective. In: *Proc. IMC '09* (pp. 35–48). New York, NY: ACM.
- Tan, P., M. Steinbach and V. Kumar (2005): *Introduction to Data Mining*. Boston, MA: Addison-Wesley Longman Publishing.
- Teevan, J., D. Ramage and M. Ringel Morris (2011): #TwitterSearch: A Comparison of Microblog Search and Web Search. In: *Proc. WSDM '11* (pp. 35–44). New York, NY: ACM.
- Wang, B., C. Wang, J. Bu, C. Chen, W. V. Zhang, D. Cai, and X. He (2013): Whom to Mention: Expand the Diffusion of Tweets by @ Recommendation on Micro-Blogging Systems. In: *Proc. WWW '13* (pp. 1331–1340). New York, NY: ACM.

- Wang, G., T. Konolige, C. Wilson, X. Wang, H. Zheng and B. Y. Zhao (2013): You are How you Click: Clickstream Analysis for Sybil Detection. In: *22nd USENIX Security Symposium, August 14–16, 2013, Washington, D.C.* (pp. 241–256). Berkeley, CA: USENIX Association.
- Wang, G., X. Zhang, S. Tang, H. Zheng and B. Y. Zhao (2016): Unsupervised Clickstream Clustering for User Behavior Analysis. In: *Proc. CHI '16* (pp. 225–236). New York, NY: ACM.
- Wen, X., Y. Lin, C. Trattner and D. Parra (2014): Twitter in Academic Conferences: Usage, Networking and Participation over Time. In: *Proc. HT '14* (pp. 285–290). New York, NY: ACM.

Panels

Significance and Relevance of Information Science in German-language Countries

A Panel Discussion Devoted to the 65th Birthday of Wolf Rauch

*Wolf Rauch*¹, *Rainer Kuhlen*², *Wolfgang G. Stock*³, *Christian Wolff*⁴, *Christa Womser-Hacker*⁵, *Christian Schlögl*¹

University of Graz, Austria
wolf.rauch@uni-graz.at

University of Konstanz, Germany
Rainer.Kuhlen@uni-konstanz.de

Heinrich Heine University Düsseldorf, Germany
stock@phil.hhu.de

University of Regensburg, Germany
christian.wolff@uni-regensburg.de

University of Hildesheim, Germany
womser@uni-hildesheim.de

Keywords: information science; German-language countries; review; importance

1 Introduction

30 years ago, in his inaugural address on the occasion of the foundation of the Institute of Information Science at the University of Graz, Wolf Rauch

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 252–263.

talked about significance and research topics of information science (Rauch, 1988). Two key statements, among others, were:

- Information science is at least as important as computer science (ibid.: 6).
- In the U.S. and in England information science is a small but well established discipline. This is in contrast to German-language countries, where it was much difficult for information science to gain ground (ibid.: 14 f.).

While the first statement was probably a little bit provocative, the second one was based on a realistic assessment of the status of academic information science at that time.

On the occasion of the 65th birthday of Wolf Rauch, the panel aims at performing a critical review of the actual situation of information science. In particular, the following questions will be discussed in more detail:

1. How is the actual role of information science in German-language countries?
2. What are possible reasons for the changed role of information science today? Did the research topics of information science considerably change in the last decades? Are there still research topics which can be primarily attributed to information science nowadays?
3. What can information science in German-language countries do to increase its role in the future?

2 Statements by panelists

2.1 Statements by Wolf Rauch

Question 1

Information science is the big success story of our time. The topic of information science (impact of computer and telecommunication in society, economy and science) is nowadays the most important development in all parts of the world. The science and technology behind this development, like information retrieval, language processing, artificial intelligence, data mining etc. are core elements of information science.

However, this development is concentrated in the USA and Asia, not in Europe, not in the German speaking countries. And it is mainly focused in private companies (Google, Apple, Amazon, etc.) not at universities.

Therefore it is no surprise that information science as a scientific discipline plays, in the German speaking countries, a minor role in science and economy.

Question 2

As already mentioned above, it is the enormous economic and social impact of information and communication technology which makes information science so important. Information science is the driving force behind this development, focusing on the traditional research topics of information science, like information retrieval, data mining, man-machine interface, artificial intelligence etc. This development is driven by the private sector.

In the public sector, security questions like cybercrime, cyberwar and the protection of critical infrastructure are in the center of interest. Here intelligence agencies of the military and the police are the driving forces, private security companies gain importance. Again universities play a minor role.

Question 3

What is missing in the new development is the consideration of the civil society. Privacy, information ethics, intellectual property; the role of the citizen versus private companies and governments; the long-term, unintended and indirect effects of an information society (information assessment); democracy is in danger, “Big Brother” and the “Transparent Citizen” are already reality.

These topics should be in the center for information science in Europe, mainly in the German speaking countries. This discipline of science is not in the focus of private companies, and only of marginal importance for the public security sector. Here although information science has a long tradition at European universities. In these questions Europe is more alert than the USA or the Asian countries. Let’s concentrate on this niche which will be a key element for the development of future information societies.

2.2. Statements by Wolfgang G. Stock

Core topics and current enhancements of information science (question 2)

There is no doubt on the core topics of information science. It is the scientific discipline which describes and analyzes the pragmatics and dynamics of information (Rauch, 1994, 2004). In contrast to data, which are the objects of computer science, information science deals with knowledge and – while transmitting knowledge – information. For Kuhlen (1990: 14), information is “knowledge in action” and so it is always dynamic. “Information is operative action-relevant knowledge”, Rauch (1988: 26) correctly states. Henrichs (2014: 193) defines information science as the science of knowledge organization, more precisely as “the science of the application and utilization of knowledge.” As knowledge can be seen as the content of information, information science is the science of information content (Stock & Stock, 2013: 8). Its core sub-disciplines (Stock & Stock, 2012) include knowledge representation and information retrieval (Stock & Stock, 2013), informetrics, as well as application-oriented subjects such as knowledge management, information literacy, and the information markets (Linde & Stock, 2011). As knowledge and information play outstanding roles in the currently emerging knowledge society, information science is – next to computer science – one of the fundamental sciences of the societies of the 21st century.

With the upcoming knowledge society, new research topics in information science emerged. We would like to introduce some examples. (1) Information systems of all kinds entered the work place and the people’s private lives. It is important to analyze and to evaluate such information systems. It is a task of information science to establish evaluation models and methods which exceed disciplinary boundaries and unite approaches from, e.g., information systems research, business administration and marketing in order to cover all critical aspects of information systems (cf., e.g., Schumann & Stock, 2014). (2) The information behavior of users changed. Some years ago, information science described information search behavior (say, on Web search engines and professional information services). Nowadays, users additionally exhibit information production behavior as posting texts, images and videos on Social Network Services (SNSs) and information reception behavior (reading and reacting to such posts). Social media, e.g. asynchronous SNSs as Facebook (Knautz & Baran, 2016) or synchronous SNSs as YouNow and Periscope (Scheibe, Fietkiewicz & Stock, 2016) form new research topics. Therefore, all kinds of information behavior became object of

information science. (3) In knowledge societies, a new type of cities, namely the so-called “informational” or “smart” city came into reality (Stock, 2011). In cooperation with other disciplines information science plays a central role in “informational urbanism” (Barth et al., 2017), which analyzes prototypical cities of the knowledge society and gives hints to construct new cities (as Songdo in South Korea) or to reconstruct old cities (as, e.g., Vienna, Oulu or Barcelona) in terms of “informativeness” or “smartness.” These three examples impressively demonstrate the significance of information science in the knowledge society.

Decision Delphi: Information science in German-speaking countries or not (questions 1 and 3)

In the German-speaking countries, information science never cleared of the hurdle of the critical mass and carved out a more or less miserable existence as an exotic “Orchideenfach” at the universities in Germany, Austria and Switzerland. According to Henrichs (2014: 214), such a small scientific discipline has no prospects in higher education. As information science institutes in Konstanz and in Saarbrücken are already closed and the Düsseldorf institute is to be shut down in 2022, Henrichs’ prognosis seems to be confirmed. Is there any solution of this unfortunate development?

Decades ago, Rauch (1979) introduced the decision Delphi. In contrast to classical Delphi (forecasting situations) and policy Delphi (analyzing a situation), decision Delphi prepares, assists and makes decisions. With a decision Delphi reality will be created. Precondition is to introduce crucial decision makers into the Delphi study: “In a decision Delphi reality is not predicted or described; it is made” (ibid.: 163). For Rauch, the case study was the system of scientific information and documentation in Austria (Rauch, 1978), for us it would be the system of researching and teaching information science in German-speaking languages. We have to bring together information scientists, policy makers in national parliaments, policy makers in higher education and players on information markets in order to create an information science which the German-speaking countries really need to master the challenges of knowledge society.

2.3 Statements by Christian Wolff

Question 1

The situation of information science in German speaking countries (D-A-CH) is ambivalent in my view: On the one hand, topics in information science have gained importance recently. One might think of digital society as a very broad phenomenon, of usage of information and information systems in everyday life (information behavior), or the general trend towards more explicitness and more measurability in science in general. These topics are typical information science research domains. On the other hand, we observe that topics previously associated with information science are today being occupied by computer science and related fields (software engineering, management information systems). One example could be human-computer interaction, which decades ago was more established in information science than in typical computer science programs. With the advent of media informatics as a novel branch of applied computer science this field has been successfully integrated in computer science programs.¹

Another aspect of this ambivalence is that today we have (arguably) more information science students than ever before and a reasonably strong position of the field at universities of applied science, especially at those who offer LIS programs for librarians. It might even be the case that more information science than actual library science professors teach in this context. At the same time, the situation at universities appears to be less promising: In Germany, several of the original information science departments have been closed (FU Berlin, Saarbrücken, arguably Constance as well) or will be closed in the next years (Düsseldorf). The rest is flourishing (Berlin, Hildesheim, and Regensburg). The future of information science at universities in Austria is unclear, in Switzerland there is no comparable degree program at a university.

This does certainly not mean that there is no information science research in D-A-CH, quite the opposite is the fact: There is a D-A-CH information science conference (i.e., this conference) which has recently become more international and its proceedings are published in English. There are academic societies for information science, and ties with the international com-

¹ In 2016, the German Gesellschaft für Informatik made HCI part of their general curriculum recommendations for computer science for the first time (Gesellschaft für Informatik, 2016).

munity have become stronger in the last years. In 2016 the ASIST European chapter was awarded the chapter of the year award by ASIST and it has a strong participation from German-speaking countries. Only recently, information scientists from German-speaking countries have been involved in top positions (e.g., as program co-chair) in the organization of the first ever ASIST meeting to be held in Europe (79th ASIST Annual Meeting, Copenhagen).

What is missing, is (strong) political awareness of the field and the willingness to strengthen information science. When politicians talk about the digital society, they almost certainly do not think of information science (rather of computer science or artificial intelligence).

In the past there has been that *kairos* in the aftermath of the Sputnik shock when western societies saw an urgent need to strengthen their information infrastructure which among other things led to the introduction of information science as an academic curriculum (Herner, 1984: 159). We do not have a comparable opportunity right now.²

Question 2

As mentioned above, some of the traditional topics of information science have been picked up by computer science (informatics) as well. This is obvious for HCI as well as for information retrieval (IR), one of the traditional core topics in information science. Here, it has to be admitted that computer science traditionally played a large role concerning the technical dimension of IR. With user studies and empirical work becoming more commonplace in computer science due to the larger role of human-computer interaction in media informatics, for example, the breadth of IR studies in computer science contexts has increased as well.³

2 Or do we? The strategy to fill top positions in information infrastructure institutions with scientists (professors) as demanded by the national council on science (Wissenschaftsrat) is a huge chance for information science which has so far only partially been made use of, though (due to various reasons, among them the small number of qualified and willing academic information science personnel).

3 Personal anecdote: Immediately after my Ph. D. in information science I started to work in a traditional CS institute. The (well-renowned) database professor there learned of my work in IR and virtually laughed at me, informing me about the marginal importance of IR compared with the database field. A few years later, the guy had successfully starting working in the field of search engine technology ...

Still, it can be assumed that all work that has a methodological basis in the social sciences – from scientometrics and bibliometrics to information behavior and information interaction in everyday life – will not be fully embraced by computer science and remain a stronghold of information science work. It should be seen as one of the core elements of a genuine information science identity. While this argument is driven by methodological and subject-related arguments, it does not guarantee a future strengthening of information science.

Question 3

Visibility, especially international visibility is an important aspect. Regarding this, it is a surprise how little information scientists from German-speaking countries have been present in international information science conferences and journals.⁴ Certainly there are exceptions, and if one starts to wonder why information scientists have not taken part more actively in ASIST activities, the fact that ASIST has been the national information science society of the United States until 2013 should be recalled. Things are obviously changing right now. The general trend towards internationalization has arrived quite late for information science: International networking has gained importance and is more visible.

Some parts of the information science community have been active in neighboring fields of computer science, especially in media informatics (I am speaking pro domo, of course). While this has been quite successful in the sense that several information scientists have become computer scientists or (media) informatics professors. At the same time there is a tendency to fill information science positions with computer scientists as can be observed in information infrastructure institutions. For a larger community this bidirectional exchange might be fruitful and inspiring. For a community as small as information science, influx from neighboring fields can quickly be seen and felt as dangerous.

The question arises whether information science shouldn't focus more on its unique selling propositions like information behavior, research data management, information literacy, bibliometrics, information retrieval and everything connected with information infrastructure in general.

⁴ Christian Schlögl has published some bibliometric analyses that touch this point, so there is no need to go into further detail here (cf. Schlögl, 2013).

Summing up, there are plenty of exciting research challenges for information science, which can be used for sharpening the scientific profile of a unique discipline. At the same time, its degree of institutional establishment is quite lackluster.

What is surprising in this context is the fact that the very visible and important field of libraries is even less established as an academic field while at the same time it is clearer than in the case of information science that no single replacement field exists, certainly not computer science. From this observation and a positive judgement of the importance of the LIS field for society (information literacy and behavior, dealing with the post factual, giving orientation in the digital society) I draw the conclusion that there is a rising need for this field. Unfortunately, we have not found the political lever for adequate institutional consequences from this view yet.

Addendum

Last October at the ASIST conference in Copenhagen, Hazel Hall (Edinburgh) presented a study of the actual size of the UK information professional labor market (including librarians) (Hall & Raeside, 2016). The quite impressive number of 86.000 positions in the UK was observed. I do not know of a comparable study for German speaking countries. It might be quite rewarding to produce similar data here – with results stating that there are well beyond 100.000 DACH information professionals?

2.4 Statements by Christa Womser-Hacker

Question 1

In my view information science should have an important role since it delivers the scientific basis for the information society. Many information-related issues are relevant for recent developments. Information science is a small discipline which cannot compare itself to other disciplines such as computer science. I.e. that information science cannot defend a single position. We cannot keep the borders of disciplines; we should / are forced to form coalitions.

Question 2

Information science does not have a binding and mandatory curriculum for research and teaching. On the one hand, this is positive because it opens opportunities for dynamic adaptation and openness for changes within technol-

ogy, society and human behavior. On the other hand, teaching and research approaches tend to be very individual and often appear as isolated from each other. Information science could play an important role representing the interface between the big sciences such as computer science and the Humanities.

Information science has relevant and current topics and is flexible to move into new directions. Currently, there are no topics, e.g. information retrieval, that are exclusively dealt with within information science but we should demonstrate the benefit which information science can bring in.

Question 3

Try to get to know each other better and work on overcoming heterogeneity. My perspective is to find partners at the European level. I recently experienced very intensive and excellent collaborations with colleagues from the Scandinavian countries and from Southern and South-Eastern Europe.

References

- Barth, J., K. J. Fietkiewicz, J. Gremm, S. Hartmann, A. Ilhan, A. Mainka, C. Meschede, and W. G. Stock (2017): Informational urbanism. A conceptual framework of smart cities. In: *Proceedings of the 50th Hawaii International Conference on System Sciences, January 4–7, 2017, Waikoloa Village* (pp. 2814–2823). IEEE Computer Society.
- Gesellschaft für Informatik (2016): Empfehlungen für Bachelor- und Masterprogramme im Studienfach Informatik an Hochschulen. July 2016. Gesellschaft für Informatik e.V. https://www.gi.de/fileadmin/redaktion/empfehlungen/GI-Empfehlungen_Bachelor-Master-Informatik2016.pdf <10.1.2017>
- Hall, H., and R. Raeside (2016): Mapping the UK information workforce in the library, archives, records, information management, knowledge management and related professions. In: *Proceedings of the 79th ASIS&T Annual Meeting: Creating Knowledge, Enhancing Lives through Information & Technology, Copenhagen, Denmark*.
- Henrichs, N. (2014): *Menschsein im Informationszeitalter. Informationswissenschaft mit Leidenschaft und missionarischem Eifer*. Hülsbusch.
- Herner, S. (1984): Brief history of information science. In: *Journal of the American Society for Information Science*, 35 (3), 157–163. [doi:10.1002/asi.4630350308](https://doi.org/10.1002/asi.4630350308)

- Knautz, K., and K. S. Baran (Eds.) (2016): *Facets of Facebook. Use and Users*. De Gruyter Saur.
- Kuhlen, R. (1990): Zum Stand pragmatischer Forschung in der Informationswissenschaft. In: J. Herget & R. Kuhlen (Eds.): *Pragmatische Aspekte beim Entwurf und Betrieb von Informationssystemen* (pp. 13–18). UVK.
- Linde, F., and W. G. Stock (2011): *Information Markets. A Strategic Guideline for the I-Commerce*. De Gruyter Saur.
- Rauch, W. (1978): Republik Österreich: Ergebnisse einer Delphi-Studie über den gegenwärtigen Stand und zukünftige Entwicklungen des wissenschaftlich-technischen Informations- und Dokumentationswesens in Österreich. In: W. Rauch & G. Wersig (Eds.): *Delphi Prognose in Information und Dokumentation* (pp. 115–275). Saur.
- Rauch, W. (1979): The decision Delphi. In: *Technological Forecasting and Social Change*, 15, 159–169.
- Rauch, W. (1988): *Was ist Informationswissenschaft? Akademische Antrittsvorlesung* (Grazer Universitätsreden; 32). Kienreich.
- Rauch, W. (1994): Informationsdynamik und Informationspragmatik. In: W. Rauch et al. (Eds.): *Mehrwert von Information – Professionalisierung der Informationsarbeit*. Proceedings des 4. Internationalen Symposiums für Informationswissenschaft (pp. 15–18). UVK.
- Rauch, W. (2004): Die Dynamisierung des Informationsbegriffs. In: R. Hammwöhner et al. (Eds.): *Wissen in Aktion. Der Primat der Pragmatik als Motto der Konstanzer Informationswissenschaft. Festschrift für Rainer Kuhlen* (pp. 109–117). UVK.
- Scheibe, K., K. J. Fietkiewicz, and W. G. Stock (2016): Information behavior on social live streaming services. In: *Journal of Information Science Theory and Practice*, 4 (2), 6–20.
- Schlögl, C. (2013): International visibility of European and in particular German-language publications in library and information science. In: Hans-Christoph Hohbohm (Ed.): *Informationswissenschaft zwischen virtueller Infrastruktur und materiellen Lebenswelten, Potsdam, 19th–22nd March 2013*. Proceedings of the 13th International Symposium on Information Science (ISI 2013) (pp. 50–62). Hülsbusch.
- Schumann, L., and W. G. Stock (2014): The Information Service Evaluation (ISE) model. In: *Webology*, 11 (1), Article No. 115.
- Stock, W. G. (2011): Informational cities. Analysis and construction of cities in the knowledge society. In: *Journal of the American Society for Information Science and Technology*, 62 (5), 963–986.

Stock, M., and W. G. Stock (2012): Was ist Informationswissenschaft? In: O. Petrovic et al. (Eds.): *Informationswissenschaft. Begegnungen mit Wolf Rauch* (pp. 389–407). Böhlau.

Stock, W. G., and M. Stock (2013): *Handbook of Information Science*. De Gruyter Saur.

“Information and Learning”

Research at the Interface between Information Science and the Learning Sciences

Antje Michel

Fachhochschule Potsdam, Potsdam, Germany
michel@fh-potsdam.de

Abstract

What role does the form of information play in the way we learn? How relevant are the specific information-related behavioural patterns of different social groups or “knowledge cultures” to the didactic design of teaching and learning processes? The idea of this panel is to provide an opportunity to explore and discuss different ideas on the relevance of “Information” as the core category of information science and of research perspectives of information science to the context of teaching and learning processes. The goal is to encourage discourse and networking between scholars working at the hitherto neglected interface between information science and the learning sciences.

Keywords: information; learning; information didactics; learning sciences; information literacy; information-related behaviour; information behaviour research; information sociology; epistemic cultures

1 Introduction

While the analysis of information behaviour is a well-established research field in information science, the application of this research to the analysis and development of didactical processes is still fairly unusual within the dis-

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 264–268.

cipline (Hobohm, 2015). But with the emerging relevance of the fostering of lifelong learning skills both within the educational process and (as a reaction to this) in the field of information literacy studies, this focus is developing into a matter of crucial importance within the information sciences. Alongside the increasing complexity of the information landscape as a result of digitalization and globalization there has been a corresponding increase in the relevance of the implementation of information literacy as part of teaching and learning processes (Framework for Information Literacy for Higher Education, 2015).

Yet the transfer of knowledge between the information and learning sciences is rudimentary. “Information” has not yet been established as an independent category within existing didactic concepts and models (Ballod, 2007). Instead, concepts and methods for teaching and learning processes are based on the *content* to be transferred or (at least in Germany) on the *media form* (Kron, Jürgens & Standop, 2014: 25).

However, in order to benefit from teaching and learning methods beyond existing disciplinary boundaries, it is worth pursuing the analysis of the particular *forms of information* as well as the *practice of obtaining this information* within the different “epistemic cultures” (Knorr Cetina, 2002) for two reasons (Michel, 2016).

First, in order to be aware of discipline-specific differences within teaching and learning processes. Familiarity with the way each discipline typically approaches these concepts is fundamental to the design of study courses. However this particularly applies to those aiming to develop successful advisory and teaching resources in professional fields related to information practice since practitioners in these information-related fields do not usually share the same knowledge culture as their customers.

Second, to explore similarities which could be used not only to adapt and extend teaching and learning methods beyond existing disciplinary boundaries but also to develop interdisciplinary learning strategies. The increasing importance of complex interdisciplinary relationships between fields in our high tech environment (Langemeyer, 2015) underlines the urgent need for the further development of interdisciplinary learning strategies. This clearly indicates how much the learning sciences could profit from an analysis of the category “information” in relation to teaching and learning processes.

But how relevant is an analysis of the role played by information and “obtaining information” within learning processes for the information sciences themselves? First, this kind of research at the threshold of the information

sciences and learning sciences could potentially be further developed to provide adequate didactic models for practical information literacy teaching and learning processes, as stated above. Second, by extending the fundamental relevance of “information” (as a core category of the information sciences) to teaching and learning processes, a contribution could be made to the interdisciplinary profile of the information sciences. And thirdly, the very act of describing what is considered as “information” in the various knowledge culture branches and how the practice of “obtaining information” is organised offers empirical access to the widely discussed definition of the category “information” (cf. for example Hjørland, 2015).

2 Aim of the ad hoc panel

The goal of this panel is to encourage discourse and networking between scholars working at the hitherto neglected interface between information science and the learning sciences. This is the reason the event has been organised as an ad hoc panel. It means that there is as yet no definitive list of contributors. Scholars wishing to introduce and discuss their fields of interest in the form of a short presentation (approx. 15 min.) are invited to contact Antje Michel in advance (michel@fh-potsdam.de).

3 Panel organisation

1. *Introduction* (20 minutes)

Antje Michel will present the thematic focus of the workshop, explain general organisational details relating to the panel and introduce the participants. To make sure the audience’s core areas of interest and expertise are fully integrated into the discussions a survey will be conducted during the introduction using an electronic voting system.

2. *Panellist Presentations* (50 minutes)

Each panellist will present his or her approach to the thematic focus of the panel in a short talk of 10–15 minutes. Immediately following the

talks there will be time for the audience to ask any short questions they may have on points of comprehension.

3. *Concluding discussion between panellists and audience (20 minutes)*

The impulses generated by the presentations and the electronic voting results will be incorporated into a concluding discussion between all participants, in which the audience is invited to participate.

4 Potential thematic focus of the panel

The potential topics have been deliberately formulated as broadly as possible so as to maximise the accessibility of the new focus of the threshold between information sciences and learning sciences to existing potential research areas. The panel is an opportunity both to present completed or current research projects pertaining to the thematic focus and to define areas urgently needing research or develop new theories.

Potential fields of interest could include, but are not limited to:

- the role played by information search methods, processing and dissemination in learning and teaching processes
- didactic models and their use in the teaching of information literacy
- information behaviour patterns in different epistemic cultures and their relevance for teaching processes
- the role of information behaviour patterns in the development of (lifelong) learning skills

5 Panel chair

Antje Michel is Professor of Information Didactics and Knowledge Transfer at Potsdam University of Applied Sciences (FHP). The main focus of her research and teaching activity at FHP is on developing the new research field of information didactics, combined with the interdisciplinary integration of the related concepts and research results into academic teaching and knowledge transfer. She studied sociology at Göttingen University and the Freie

Universität in Berlin and was awarded a doctorate in sociology by the Freie Universität Berlin in 2003. After gaining a postgraduate qualification in academic librarianship she held posts at the Max Planck Digital Library and Munich University Library.

References

- Ballod, M. (2007): *Informationsökonomie – Informationsdidaktik. Strategien zur gesellschaftlichen, organisationalen und individuellen Informationsbewältigung und Wissensvermittlung*. Bertelsmann.
- Framework for Information Literacy for Higher Education (2015). Chicago. <http://www.ala.org/acrl/standards/ilframework>
- Hjørland, B. (2015): The concept of information – again and again. In: *Cybernetics & Human Knowing*, 22 (1), 67–69, 78–80.
- Hobohm, H.-C. (2015): Informationsdidaktik + Informationsverhaltensforschung = Informationskompetenz. Eine Gleichung mit drei Unbekannten. In: A.-K. Mayer (Ed.): *Informationskompetenz im Hochschulkontext. Interdisziplinäre Forschungsperspektiven*. Pabst Science Publ (pp. 29–42).
- Knorr Cetina, K. (2002): *Wissenskulturen. Ein Vergleich naturwissenschaftlicher Wissensformen*. Suhrkamp.
- Kron, F. W., E. Jürgens and J. Standop (2014): *Grundwissen Didaktik*. UTB.
- Langemeyer, I. (2015): *Das Wissen der Achtsamkeit. Kooperative Kompetenz in komplexen Arbeitsprozessen*. Waxmann.
- Michel, A. (2016): Informationsdidaktik – Skizze eines neuen informationsdidaktischen Forschungsfelds. In: *Information – Wissenschaft & Praxis*, 67 (5–6), 325–330.

Examining Research Integrity¹

Michael Seadle

Berlin School of Library and Information Science, Germany

seadle@ibi.hu-berlin.de

Abstract

Research integrity issues fill the academic news, and include plagiarism, data falsification and image manipulation. Integrity violations are complex because of the gray zones between where bad practice ends and genuine malpractice begin. No real consensus exists about the boundaries, even though many people have strong opinions. The goal of this panel is to engage in a scholarly discussion about integrity issues using specific examples drawn from the book “Quantifying Research Integrity” (Seadle, 2017).

Keywords: research integrity; plagiarism; data falsification; image manipulation

Proposal

Research integrity issues fill the academic news. The blog “Retraction Watch” tries to record each retraction in major scholarly journals, especially in the natural and medical sciences.² In Germany more and more dissertations come under formal scrutiny because of plagiarism (Fischer, 2016). The US Office of Research Integrity has focused especially on tools to uncover

1 The author wishes to thank the HEADT Centre [Humboldt Elsevier Advanced Data and Text Centre] for finding the research in this proposal

2 <http://retractionwatch.com/>

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 269–271.

image manipulation.³ Diederik Stapel in the Netherlands admitted that he made up data and gave those data to doctoral students to use in their doctoral dissertations (Budd, 2013). Scandals abound, but not all accusations are true and proof can be difficult to establish. Opinion can also shift over time, as shown in the case of Cyril Burt, who was posthumously accused of falsifying data until Robert Joynson (1989) found inconsistencies in the case for the prosecution.

Integrity violations are complex because of the gray zones between where bad practice ends and genuine malpractice begin. No real consensus exists about the boundaries, even though many people have strong opinions. When, for example, does standardized language used to express a statistical result in a social article become plagiarism? Computerized matching systems will generally flag all overlaps in expression and will add them to a score that triggers an investigation. Outliers can also be a problem in statistical tests. Throwing out some outliers is legitimate, but throwing out too many, even for good reasons, may begin to look like falsification. Cleaning up an image display from a scientific test may only make the results clearer or the manipulation may render the image inauthentic because it exaggerates the results.

The goal of this panel is to engage in a scholarly discussion of these issues like these using specific examples drawn from the book *Quantifying Research Integrity* (Seadle, 2017).

An issue of concern in the scholarly community is how many integrity violations remain to be found, and how seriously do they affect the reliability of research. Replication studies can help (Camera et al., 2016). But many kinds of research are difficult to replicate because details in the descriptions are missing and because circumstances change the database from when the original tests were made (Call et al., 2016). Scholarship, especially the natural sciences, are supposed to build on past results in the expectation that the results are reliable, but applying contemporary standards to past results may give false results, since the standards change over time. Image manipulation is a good example, because the first innocent but careless uses of Photoshop have become associated with deliberate manipulations that change the nature of the image. Plagiarism is the easiest form of integrity violation to uncover, because computing systems today can easily compare texts, but many jour-

3 Office of Research Integrity [2016], "Forensic Tools". <http://ori.hhs.gov/forensic-tools>

nals fail to use the standard tools for cost reasons. Much research remains to be done.

A panel discussion similar to this proposal took place at the 2014 iConference in Berlin and produced lively debate in a full room. The proposed members for this panel include people who either have experience analyzing research integrity issues, or have positions where they must deal with the results of integrity problems.

References

- Budd, John (2013): The stapel case: An object lesson in research integrity and its lapses. In: *Synesis: A Journal of Science, Technology, Ethics, and Policy*, 4 (1): G47–G53.
- Call, Andrew, Max Hewitt, Jessica Watkins, and Teri Lombardi Yohn (2016): Changes in the i/b/e/s database and their effect on the observed properties of analyst forecasts, 2016-06-03. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2788140
- Camerer, Colin F, Anna Dreber, Eskil Forsell, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Johan Almenberg, Adam Altmejd, Taizan Chan, and Others (2016): Evaluating replicability of laboratory experiments in economics. In: *Science*, 351 (6280): 1433–1436.
- Fischer, Benjamin (2016): Uni: Backhaus darf Dokortitel behalten. In: *Ostsee Zeitung*, 11.11.2016. <http://m.ostsee-zeitung.de/Nachrichten/MV-aktuell/Politik/Uni-Backhaus-darf-Dokortitel-behalten>
- Joynson, Robert B. (1989): *The Burt affair*. Taylor & Frances/Routledge.
- Seadle, Michael (2017): *Quantifying Research Integrity*. Morgan Claypool. [doi:10.2200/S00743ED1V01Y201611ICR053](https://doi.org/10.2200/S00743ED1V01Y201611ICR053)

**Gerhard Lustig Award
Papers**

Data Strategy Model

A Reference Model to Develop Data Strategies

Fabian Muster

HTW Chur, Switzerland
fabian.muster@bluewin.ch

Keywords: data strategy; reference model; data management; information science; strategy transformation

1 Introduction and outline

Data, data, data ... never before companies have had as much data as today. And the potential to gain profitability from and with data is enormous – but often not used! Either there is no awareness or a lack of qualification within the company that makes it hard or even impossible to benefit from those data. A comprehensive data strategy can enable companies to benefit from data. But data strategies are not so easy to establish and transform – and they are not even available as a model! Therefore, the at hand reference model can help closing the gap between the “we need” and the “we have” a data strategy.

The Master Thesis answers the questions why a data strategy is relevant and why a reference model is helpful, where in the company and how a data strategy has to be established, identifies the most important components of such a strategy and creates a reference model for data strategies. Complemented with guidelines for the transformation and for a maturity model, a RACI model for roles and responsibilities and an implementation roadmap, the master thesis presents a feasible reference model for data strategies.

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 274–277.

2 Developing the reference model for data strategies

Based on a literature and market research, possible components for a data strategy were identified. Ensuing, the identified components were assessed and a justification for each of the identified components has been made to decide on the acceptance for the data strategy model.

Further, similar components were grouped and functions/features introduced to consider complementary components. The reference model also distinguishes supporting/dependent functions. But unlike the components or functions/features, they are not directly related to the components of a data strategy however they are important for the transformation.

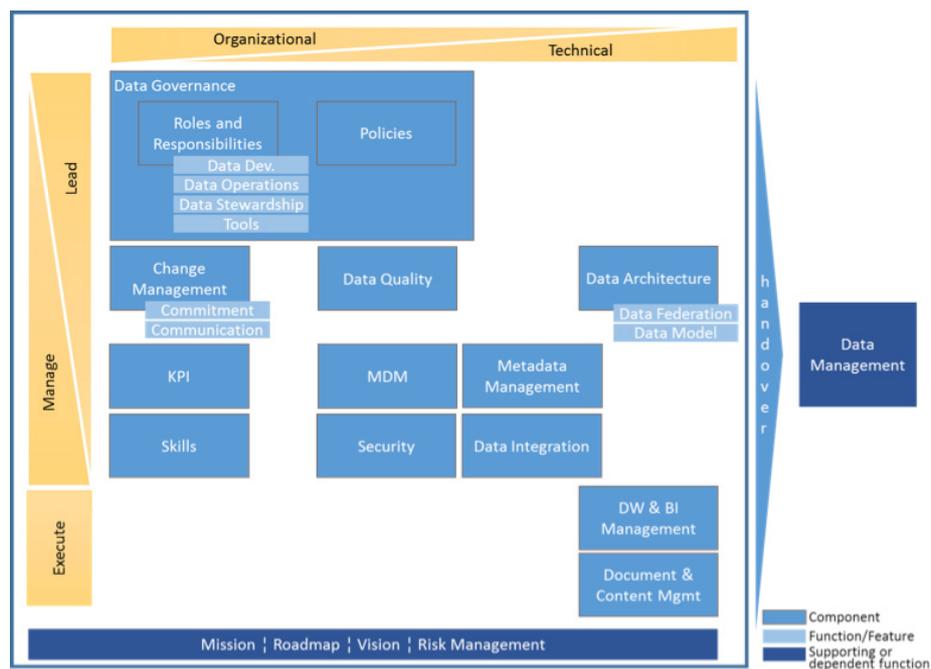


Fig. 1 Data strategy reference model

In addition, implementation levels and a differentiation between a more organizational or technical approach were introduced. This structure supports the later transformation in the company and with that the responsibilities: the implementation levels “lead”, “manage” and “execute” mainly support the

role allocation and the differentiation between organizational and technical approach facilitates the task allocation.

With the selected components and functions, the data strategy reference model has been developed:

3 Putting the data strategy into action

With the developed reference model, the basis for implementing a data strategy has been established. In order to be able to put a strategy into action, transformation guidance is necessary. The master thesis provides guidelines and templates for the following transformation steps: First, a maturity assessment of each of the components/functions has to be done. With the result – the gap between the current and the future state – it becomes clear for a company which of the components/functions need to be implemented, enhanced or even “ignored” for the moment in case they are already well established. Second, a RACI model helps companies to identify the most important roles concerning the transformation and further development of the data strategy. With the identified gaps (maturity model) and the identified persons responsible (RACI model), the transformation itself can be executed. For this step, the master thesis also provides information in the form of instructions for the strategy program including a strategic management model.

4 The role of information science

Information scientist can play a leading role in the area of data strategies since developing and transforming data strategies requires a unique mix of skills: a data strategist needs to be able to understand the technical dimension of a data strategy since he or she is responsible for a well implemented data architecture, security concept and other more technical oriented components and functions of a data strategy. Furthermore, organizational and economic skills are required to lead and manage for example responsibilities, policies and the important change management. Last but not least, strategic thinking is expected to embed the data strategy in the company and to lead it into the

future where topics like big data, AI or semantic technologies are dependent on a comprehensive data strategy.

5 Outlook

This data strategy reference model is a first work in this specific area. The field is open for more interesting research regarding the assessment and trends of identified or even new components, the positioning of the data strategy within the company, the integration of semantic technologies, big data and many other interesting topics where information scientist can make a huge contribution.

Relationship between Downloads and Citations

The Case of Two Economics Journals

Roland List

Karl-Franzens-Universität Graz, Austria
roland.list@me.com

Keywords: citations; downloads; readership data; correlation; relation; Scopus; ScienceDirect; Mendeley

For many decades citations were the main source when analyzing the research impact of science. It was the aim of this master thesis to explore if downloads and so-called readership data from Mendeley are possible alternatives to citation data and if downloads can be used to forecast the citation impact in future years. The study is based on articles of the Journal of Environmental Economics and Management (JEEM) and the Journal of Financial Economics (JFE) published from 2002 to 2011, and compares the downloads (2002 to 2011) from ScienceDirect, the citations (2002 to 2016) from Scopus and the readership data (2008 to 2015) from Mendeley.

Scopus (2016) is an abstract and citation database, which includes at the moment of writing more than 60 million journal articles. ScienceDirect (2016) is a full-text database. Mendeley (2016) is a free reference manager and an academic social network service.

The analysis of the data sources shows a difference between the document types of Scopus and ScienceDirect though both are produced by Elsevier. An additional outcome of this study is, that more than 60 per cent of the documents were downloaded in PDF format and less than 40 per cent in HTML format. According to the number of readership, students are by far the big-

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 278–280.

gest user group. A significant difference exists in the obsolescence of downloads and citations: While the download maximum usually occurs for recent articles, it takes several years until the citation maximum is reached. An additional outcome of this study is that for the analyzed journals electronic articles are usually published prior to print articles. As a consequence, it may happen that a few articles are also cited before print publication.

Figures 1 and 2 show the relationship between downloads, citations and readership. To quantify the relationship, Spearman rank correlation was calculated. The results show a medium to high correlation between citations and downloads ($r = 0.62$ for Journal of Environmental Economics and Management – JEEM, $r = 0.79$ for Journal of Financial Economics – JFE) and between readership frequencies and citations ($r = 0.67$ for JEEM, $r = 0.66$ for JFE). The correlation between readership and downloads is medium sized and clearly lower ($r = 0.56$ for JEEM, $r = 0.58$ for JFE)

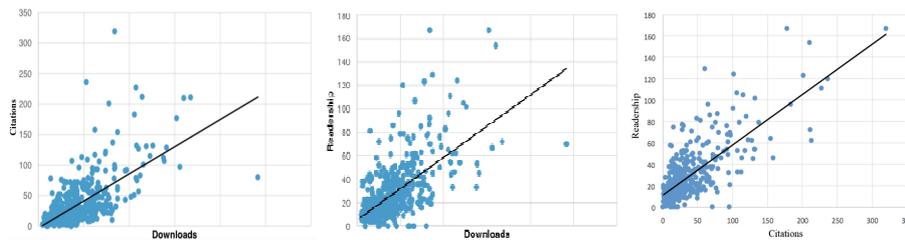


Fig. 1 Downloads vs. citations vs. readership, scattergram for “Journal of Environmental Economics and Management”, publication year (2002–2011), full length article ($n = 509$)

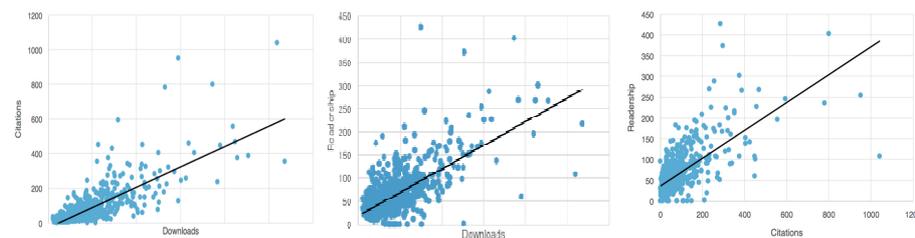


Fig. 2 Downloads vs. citations vs. readership, scattergram for “Journal of Financial Management”, publication year (2002–2011), full length article ($n = 868$)

In order to check a possible predictability of citations, I replicated an approach which was originally proposed by Kurtz et al. (2005). The results of this analysis indicate that the journal downloads can be used as a forecast for citations. Looking at all download classes in table 1 and identifying the most

likely number of citations for each shows that an increase of the download class usually goes along with an increase of the citation class. It follows that downloads can be used as a rough approximation for (future) citations. However, this does not mean that downloads can be a substitute for citations.

Table 1: Downloads (2002–2011) vs. citations (2002–2016) for FLAs published from 2002 to 2011, Journal of Financial Management (n = 870 FLAs)

Number Citations	Downloads (2002- 2011)						Total
	512-1023	1024-2047	2048-4095	4096-8191	8192-16383	>=16384	
0	1			2			3
1		1					1
2-3	7	8	1				16
4-7	10	14	8	1			33
8-15	6	62	44	6			118
16-31	7	55	118	20			200
32-63	3	17	117	62	3		202
64-127		1	52	104	14		171
128-255			6	55	26	1	88
256-511				10	16	6	32
512-1023				1	2	2	5
1024-2048						1	1
Total	34	158	346	261	61	10	870

References

- Kurtz, M. J., G. Eichorn, A. Accomazzi, C. Grant, M. Demleitner, S. Murray, N. Marimbeau, and Elwell, B. (2005): The bibliometric properties of article readership information. In: *Journal of the American Society for Information Science and Technology* 56 (2), 111–128.

Google Now and User Data

A Diary Study on Perceptions of Collection and Use of Personal Data by Google Now Cards

Alexandra S. Linhart

University of Applied Sciences Hamburg, Germany
alexandra.s.linhart@gmail.com

Abstract

In this diary study 24 participants were asked to use Google Now for 28 days, in order to learn about occurrences of negative affect caused by the use of personal data and possible factors influencing such an affect, especially situational factors and personal dispositions. The use of personal data by Google Now cards barely evoked negative affect at all. Personality, trust and situational factors proved to influence affective reactions of the participants.

Keywords: Google Now; privacy; affect heuristics; diary study

1 Introduction

Internet users do not necessarily have a carefree attitude at the use of their personal data, but privacy concern does not have a direct impact on privacy behavior (Kokolakis, 2015). Affect heuristics explain differences in attitude and behavior with the influence of emotions on risk perception. Positive affect mitigates risk perception, while negative affect intensifies it (Slovic et al., 2007). A few studies have shown that positive emotions can increase trust and willingness to disclose information online (e.g. Kehr et al., 2015). This study tries to identify factors leading to a negative affect when con-

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 281–284.

fronted with the use of personal data. Google's "personal assistant" Google Now lends itself to visualize the extent of personal data collection and use. By evaluating personal data collected with different Google products, it can present the users "cards" – small pieces of information it deems interesting for the user, e.g., news articles, traffic information for their daily commute, and the like (Guha et al., 2015).

2 Methodology

To identify possible factors influencing emotional reactions 24 participants, who had not used Google Now before, were asked to fill a questionnaire that measured their personality, the level of their general privacy concerns and their trust in Google. This was followed by the diary period over a four-week period. The participants completed a short daily online questionnaire asking them for their emotional reactions to individual Google Now cards and rate the usefulness of the cards. In case a negative affect was reported, the online questionnaire also asked for possible reasons. After the diary period, each participant took part in a personal guided interview to learn more about their experience on Google Now. Emotional reactions were collected with a semantic differential scale. To measure personality we adopted a Big-Five model scale (Rammstedt et al., 2013). As trust and general privacy concerns might be influenced by the use of Google Now, both were measured before and after the diary period. To evaluate the interrelations between measured factors, we used Spearman-Rho correlation, as most of our scales were of ordinal type.

3 Results

Google Now barely evoked negative affect. Of 1,141 card ratings collected with the diary, only 96 reported negative affect. Only in 33 cases the use of personal data was the cause of negative affect. Negative affect was most often triggered by the content Google Now presented to the users, such as news on terror attacks. Participants who felt negative about Google Now using

their personal data had a higher level of neuroticism in their personality – meaning they generally feel uneasy quickly. However, neuroticism was the only personality factor that showed an interrelation with affective reactions on Google Now. Situational factors, like the type of information used, proved to have an impact on emotional reactions. Cards using location information (e.g., cards informing them of traffic conditions) were most prone to evoke negative affect. The same can be said about cards drawing information from emails (e.g., cards presenting information on a booked flight). The predictive power of Google Now was another aspect that made the participants feel uneasy. On the other hand, inappropriate predictions, resulting in irrelevant cards, helped them to feel appeased, thinking Google does not know too much about them. The participants generally rated Google Now cards as rather useful and relevant. Interrelations between the rating of the usefulness and the affective reaction were weakly present in this study. Trust in Google increased after the diary period. This is probably due to the fact that the majority of cards resulted in a positive affect. General privacy concerns did not have an impact on affective reactions.

4 Conclusion

This study hints toward affect heuristic being a promising approach to learn more on privacy attitudes. The results strengthen the hypothesis that privacy perceptions are strongly based on emotion. Rather than asking hypothetical questions or measuring general attitudes towards privacy, more field studies should be undertaken, as situational factors have proven to influence perception of the use of personal data. Also finding a way to measure unconscious affective reactions will probably result in more insight into factors influencing affective reaction.

Acknowledgements

I would like to thank Prof. Dr. Dirk Lewandowski for his valuable and constructive help for this master thesis.

References

- Guha, Ramanathan, Vineet Gupta, Vivek Raghunathan and Ramakrishnan Srikant (2015): User Modeling of a Personal Assistant. In: *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM 2015)*, Shanghai, China, February 02–06, 2015. [doi:10.1145/2684822.2685309](https://doi.org/10.1145/2684822.2685309)
- Kehr, Flavius, Tobias Kowatsch, Daniel Wentzel and Elgar Fleisch (2015): Blissfully Ignorant: the Effects of General Privacy Concerns, General Institutional Trust, and Affect in the Privacy Calculus. In: *Information Systems Journal*, 25 (6), 607–635. [doi:10.1111/isj.12062](https://doi.org/10.1111/isj.12062)
- Kokolakis, Spyros (2015): Privacy Attitudes and Privacy Behaviour: A Review of Current Research on the Privacy Paradox Phenomenon. In: *Computer & Security*, Jul., 1–13. [doi:10.1016/j.cose.2015.07.002](https://doi.org/10.1016/j.cose.2015.07.002)
- Rammstedt, Beatrice, Christoph J. Kemper, Constanze Beierlein and Anastasiya Kovaleva (2013): Eine kurze Skala zur Messung der fünf Dimensionen der Persönlichkeit: Big-Five-Inventory-10 (BFI-10). In: *Methoden, Daten, Analysen*, 7 (2), 233–249. [doi:10.12758/mda.2013/013](https://doi.org/10.12758/mda.2013/013)
- Slovic, Paul, Melissa L. Finucane, Ellen Peters and Donald G. MacGregor (2007): The Affect Heuristic. In: *European Journal of Operational Research*, 177 (3), 1333–1352. [doi:10.1016/j.ejor.2005.04.006](https://doi.org/10.1016/j.ejor.2005.04.006)

Exploring Intergenerational Knowledge Sharing in Organizations

Potentials and Limitations of Information and Communication Technology (ICT)

Helene Brinken, Helena Kock

University of Hildesheim, Germany
helenebrinken@gmail.de, hmk@kockmail.de

Abstract

This thesis deals with an interdisciplinary approach of intergenerational knowledge sharing (IKS) from an information scientific perspective. As part of the KNOWISH project the research is embedded in an international cooperation of University of Hildesheim and the Åbo Academy in Turku, Finland. It aims at answering the questions: How does IKS unfold? Which systems are used in its context and how suitable are they? Two case studies have been conducted in two medium-sized German companies. At the same time, an additional study has been conducted by two Finnish students. Semi-structured interviews and a qualitative content analysis of the data considered generational aspects, knowledge sharing as well as ICT. This study finds that experience and work environment atmosphere are important influencing factors. Both, young and old employees can be knowledge carriers who share knowledge mutually. Another important outcome shows that Organizational Generations of Knowledge exist within the companies. These are defined by chronological and professional age.

Keywords: intergenerational; knowledge sharing; generation; organizations; information and communication technology; tacit knowledge

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 285–289.

1 Introduction

IKS becomes more and more crucial while industrialized countries are facing the demographic change and the potential loss of many knowledgeable employees. This master thesis aimed at answering the questions: How does IKS unfold? Which systems are used in its context and how suitable are they?

2 State of the art

Several areas of research have to be considered such as generation theories, aspects of knowledge and the state of the art of IKS.

In generation theories, three factors contribute to differences in values and attitude: The Generation Effect meaning differences in socialization, the Lifespan Effect due to private living circumstances and the Age Effect including aging processes (Klaffke, 2014). Mostly, generations are classified into Veterans, Baby Boomers, X, Y and Z differentiated by birth year (Hüppe, 2014; Klaffke, 2014). As the transition from one generation to another is gradual this research only differentiates between older and younger to mark tendencies which appeared in the data. Dealing with knowledge sharing it is important to consider different types of knowledge such as tacit (unconscious) and explicit (conscious) knowledge (Frey-Luxemburger, 2014; North, 2016). According to Kock (2016) different disciplines such as economic science and computer science provide research with widely varying focuses. However, an integrating approach is missing.

3 Methods

Case studies have been conducted in two medium-sized German companies based on ten semi-structured interviews in each case. Respectively older and younger employees were asked questions about social aspects, age issues, systems, generational aspects and workplace-related learning. A qualitative content analysis according to Mayring (2015) included the inductive creation

of categories. Reviewing and discussing the data in teamwork turned out to be particularly helpful and ensured intersubjectivity and comparability.

4 Results

Experience and work atmosphere emerged as important influencing factors as team spirit and social ties enable IKS. Young and old employees can be knowledge carriers. Being digital native young employees have experience with ICT. Due to their life experience old employees possess social competences.

According to Kuyken (2012) the approach to consider only generation by birth year is limited. Our analysis confirms this result. Other factors such as professional age and experience have to be considered. Both case studies revealed that organizational age leads to experience with organizational processes and professional age leads to expertise. These are learned through asking questions and performing daily work.

Considering these aspects, a new approach of IKS appears necessary. Therefore, we propose the hypothesis of the concept “Organizational Generation of Knowledge” (OGoK). This term means that within an organization generations divided by the nature of their knowledge coexist. On the one hand each OGoK is characterized by shared values and characteristics of its socialization. On the other hand it is characterized by specific knowledge about the professional environment they are working in based on similar professional and organizational age.

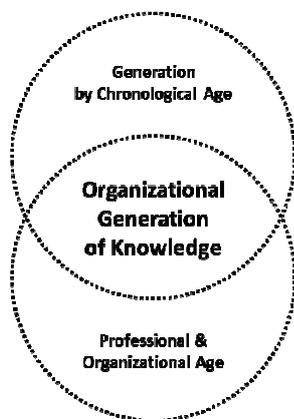


Fig. 1 Organizational Generation of Knowledge

The following example illustrates this concept: Each OGoK deals with the introduction of new social software differently.

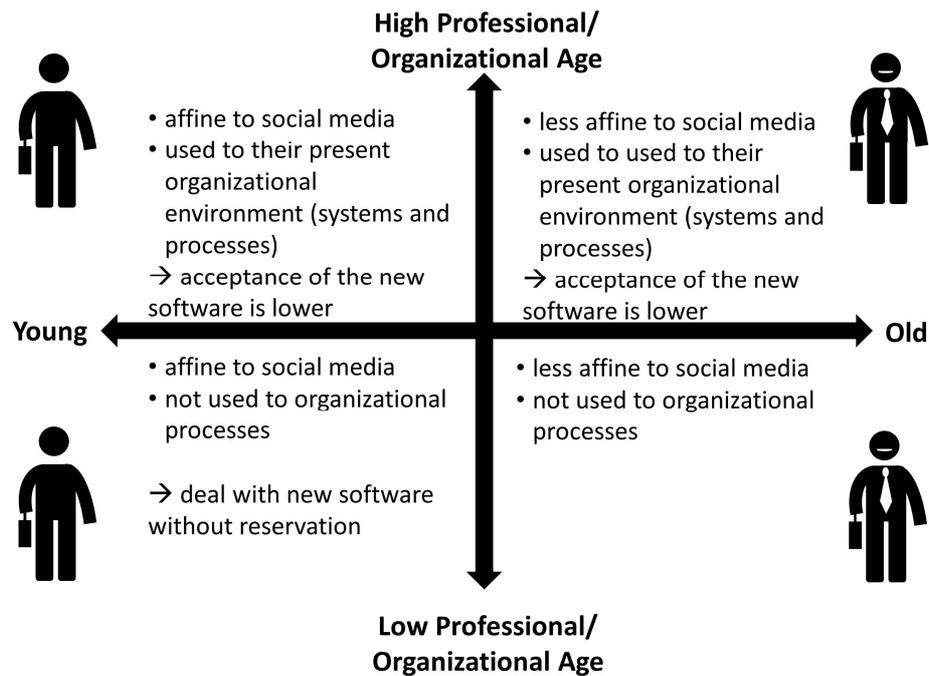


Fig. 2 Illustration – Organizational Generation of Knowledge

A variety of generations coexisting in organizations can be valuable due to different competencies of each OGoK. According to Kuyken (2012) IKS is a source of innovation. We can now specify that IKS between OGoK leads to innovation. Especially young OgoK with a low organizational age set corresponding impulses. Consequently, there is a demand for IKS between these OGoK.

In addition, the study outcomes show that the choice of ICT depends on the urgency and complexity of a problem as well as the availability of resources and knowledge carriers. As the use of ICT differs between generations, the risk of unidirectional IKS was observed.

References

- Frey-Luxemburger, Monika (2014): *Wissensmanagement – Grundlagen und Praktische Anwendung: Eine Einführung in das Management der Ressource Wissen*. 2nd ed., Wiesbaden: Springer Vieweg, Springer Fachmedien.
- Hüppe, Sebastian (2014): *Weiß ein Unternehmen, was es weiß bzw. an Wissen zu verlieren droht? Die Bedeutung von Wissensmanagement vor dem Hintergrund des demografischen Wandels*. PhD thesis, Ruhr-Universität Bochum.
- Klaffke, Martin (2014): Erfolgsfaktor Generationen-Management: Handlungsansätze für das Personalmanagement. In M. Klaffke (Ed.): *Generationen Management: Konzepte, Instrumente, Good-Practice-Ansätze*. Wiesbaden: Springer Gabler, Springer Fachmedien, pp. 3–26.
- Kock, Helena Margaretha (2016): *Intergenerational Knowledge Sharing: State of the Art*. MA thesis, Universität Hildesheim.
- Kuyken, Kerstin (2012): Knowledge Communities: Towards a Re-thinking of Intergenerational Knowledge Transfer. In: *VINE* 42 (3), 365–381.
- Mayring, Philipp (2015): *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. 12th ed., Weinheim, Basel: Beltz.
- North, Klaus (2016): *Wissensorientierte Unternehmensführung: Wissensmanagement gestalten*. 6th ed., Wiesbaden: Gabler Verlag, Springer Fachmedien.

The X(Disciplinarity) of Information Science

Sandra Balck

University of Applied Sciences Potsdam, Germany

sandra.balck@gmail.com

Abstract

The Master Thesis deals with the x-disciplinarity of information science, as it is considered as a multi-, inter- or transdisciplinary discipline in the discourse of this subject. But what in particular do these concepts mean and how are they specifically differentiated from each other? The thesis introduces the concept of disciplinarity with a historical approach and presents the development of different cooperation concepts (multi-, trans-, interdisciplinarity) as well as of the concept of transscience (Transwissenschaft). The disciplinarity of information science is analyzed on a communicational, social and cognitive level. Beside the definition of the term “information”, the paradigmatic development of information science is drafted, followed by an evaluation of relevant publications and an application of the different concepts of disciplinarity to information science. It is concluded, on the basis of a critical analysis, that real collaboration within the discipline is needed instead of blurry labels.

Keywords: history of sciences; interdisciplinarity; transdisciplinarity; multidisciplinarity; disciplinarity

1 Motivation

The concept of interdisciplinarity had been developed in the 1960s as a reaction to the increasement of scientific specialization and the differentiation of disciplines as a result of the continuing growth of knowledge. The conse-

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 290–292.

quence was an anxiety for the loss of scientific relevance; therefore interdisciplinarity had been arisen as a concept of mending. But the concept of interdisciplinarity developed rapidly into a buzz word. Other concepts – in differentiation from interdisciplinarity – emerged: multi-, pluri- and transdisciplinarity (cf. Ryser, 2016). In the upcoming discourse those concepts are summarized as x-disciplinarity. The x-disciplinary orientation of information science is indicated in numerous texts. The thesis deals with the theoretical and sociological conditions for the cooperation and networking within the disciplines. The main focus lies on information science as the “science of information” and its meaning and role in the information society.

2 Results

The disciplinary classification of information science is related to its comprehension of its main concept: information. Information science represents a paradigmatic understanding of information and dedicates itself from an action-relevant [handlungsrelevant] point of view to the use of and access to information. In order to clarify the concept of “information” x-disciplinary cooperation is needed. The paradigmatic evaluation shows, that information science has developed from a bibliographical discipline in the beginning to a pragmatic-cognitive discipline, which has mathematical and technically influences (cf. Kuhlen, 2013). But the technical side should not be left aside, because it stands in a direct relationship with recent developments: Information science as a mixture of social science and humanities with a direct influence of recent technical developments.

3 Conclusion

The analysis of relevant publications and discursive comparison (on a communicational, social and cognitive level) of specific (implicit and explicit) statements indicates that the complexity and missing disciplinary framework had been criticized a lot, especially in an international context. Often this problem is assigned to the x-disciplinarity, which prior to this was entitled as

its strength. This assumption is ascribed to a lack of questioning the concepts, which is demonstrated by insufficient terminological classifications and the lack of reflection of existing statements. So a basic desideratum is acknowledged: „Im Normalfall wird ein allgemeines, eben unterbestimmtes Verständnis von Interdisziplinarität stillschweigend vorausgesetzt oder hinsichtlich der terminologischen Regelung auf ältere Texte verwiesen, ohne dass jedoch die darin enthaltenen Regelungen kritisch geprüft würden“ (Balsiger, 2005: 137). The actual problem lays not in the cooperation, it rests in the lack of a disciplinary foundation, which is the basic concept of comprehensive cooperation. A disciplinary compartmentalization would be off target as well as the current elusiveness of the discipline. An extensive adaption of this subject needs a cooperation of different disciplines and an inclusion of the public. But it shouldn't be forgotten that „[...] keine interdisziplinäre Kompetenz [könnte], die disziplinäre Kompetenz ersetzen [...]: interdisziplinäre Kompetenz setzt disziplinäre Kompetenz voraus“ (Mittelstraß, 1987: 154). Therefore it is necessary to focus on the discipline, for creating an additional value through cooperation, instead of absorbing other concepts.

References

- Balsiger, Philipp W. (2005): *Transdisziplinarität. Systematisch-vergleichende Untersuchung disziplinenübergreifender Wissenschaftspraxis*. München: Fink.
- Kuhlen, Rainer (2013): Information – Informationswissenschaft. In: Rainer Kuhlen et al. (Eds.): *Grundlagen der praktischen Information und Dokumentation*. Berlin, Boston: Gruyter Saur (pp. 1–24.)
- Mittelstraß, Jürgen (1987): Die Stunde der Interdisziplinarität. In: Jürgen Kocka (Ed.): *Interdisziplinarität. Praxis, Herausforderung, Ideologie*. Frankfurt a. M.: Suhrkamp (pp. 152–158).
- Ryser, Vera (2016): Transdisziplinarität. Eine Bestandsaufnahme des Forschungsdiskurses. <http://blog.zhdk.ch/trans/interdisziplinaritaet/>

Interactive Pedestrian Indoor Localization

Robert Jackermeier

University of Regensburg
Universitätsstraße 31, 93053 Regensburg, Germany
robert.jackermeier@ur.de

Abstract

This thesis focuses on the improvement of pedestrian indoor localization systems through user interactions. First, a hybrid localization system for mobile devices which relies on multiple sensors is developed. It forms the basis for the design and implementation of interactions between the user and the system that aim to increase the localization accuracy. Finally, a simulation is conducted to evaluate the system's performance. It shows that localization accuracy can in fact be improved in case the user cooperates.

Keywords: indoor localization; pedestrian navigation; user interaction

1 Introduction

While GPS enables location-based services outdoors, no similar technology exists that can be used inside buildings, despite a lot of research and technological improvements in recent years. This thesis consolidates work that has been done in previous pedestrian navigation projects and introduces explicit user interaction as a new approach to improve indoor localization.

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 293–296.

2 Hybrid indoor localization

State-of-the-art pedestrian indoor localization systems rely on a multitude of sensors to estimate an accurate position and are hence called hybrid localization systems. Just like previous projects (Ebner et al., 2015; Hilsenbeck et al., 2014), the solution that was created as the first part of this thesis uses Wi-Fi fingerprinting for absolute positioning, as well as inertial sensors for relative measurements. In combination with an existing graph-based environment model (Ohm, Müller & Ludwig, 2015), the sensor information is then fused by a particle filter in order to estimate the user's position.

3 Adding user interaction

The deliberate addition of user interaction into the localization process is the main contribution of this thesis. Since the system is intended to be used as part of a pedestrian navigation application on smartphones, it can rely on the user to provide information about his whereabouts.

Figure 1 shows examples of how the user can interact with the system by entering the name of a room or selecting a nearby landmark. The interactions can be triggered by the system at various occasions during the localization process, e.g. if the uncertainty exceeds a predefined threshold.

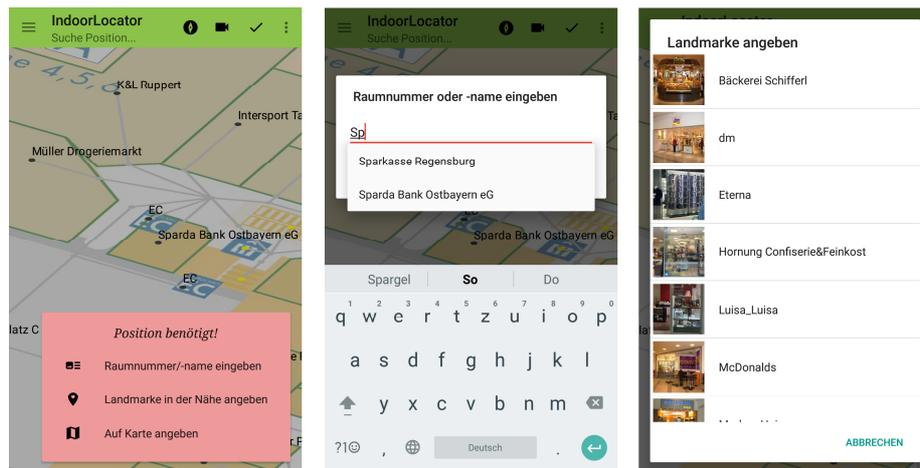


Fig. 1 The user interface of the localization system

4 Evaluation

To evaluate the system and the influence of user interactions, a simulation based on real sensor data was conducted. The data was collected in a shopping mall in Regensburg on 15 randomly selected test routes with a length between 65 and 280 meters. Various user interaction profiles were created from a combination of different factors: The input accuracy was evaluated at five levels from a theoretically perfect user to a very inaccurate user. Similarly, the absolute input error was varied. Finally, the willingness to cooperate was simulated, from an immediately cooperating to a completely uncooperative user. Based on these profiles, the following hypotheses were tested:

- H_1 : The more accurate the user input, the lower the position error.
- H_2 : The lower the absolute input error, the lower the position error.
- H_3 : The higher the cooperativeness, the lower the position error.

For every configuration, 1000 runs were simulated on each of the test routes. Localization performance was measured by the mean distance to ground truth, averaged over all runs of a configuration.

Since the data is not normally distributed and heteroscedastic, non-parametric Kruskal-Wallis tests were performed. Post-hoc Dunnett-Tukey-Kramer tests showed significant differences between every single combination of configurations on an alpha level of .05 for the first 2 hypotheses. Hypothesis 3, however, cannot be accepted completely. While it does make a difference whether the user interacts or not, the time of interaction does not significantly influence the position error in most cases.

The results also show that a good initial location is very important for accurate localization, even if no further interaction takes place. It follows that it is better to have a user interact only once accurately at the beginning, rather than multiple times, but inaccurately.

5 Conclusion and outlook

The evaluation has shown that the localization performance can indeed be improved if the user gives reasonably accurate input. However, cooperation was simply assumed here and needs to be tested in future studies.

Currently, the system is being integrated in a pedestrian navigation system, which provides more context and therefore more information to work with. In the end, even though user interaction is the focus of this thesis, it should be used sparingly in order not to overwhelm the user.

References

- Ebner, F., T. Fetzner, F. Deinzer, L. Koping and M. Grzegorzek (2015): Multi Sensor 3D Indoor Localisation. In: *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*.
- Hilsenbeck, S., D. Bobkov, G. Schroth, R. Huitl and E. Steinbach (2014): Graph-based Data Fusion of Pedometer and WiFi Measurements for Mobile Indoor Positioning. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*.
- Ohm, C., M. Müller and B. Ludwig (2015): Displaying Landmarks and the User's Surroundings in Indoor Pedestrian Navigation Systems. In: *Journal of Ambient Intelligence and Smart Environments*, 7, 635–657.

Doctoral Colloquium

**Session I – Information Literacy and Behavior,
Knowledge Management, Visualization**

Development and Evaluation of a MOOC Extension for Teaching Subject-Specific Skills in Information Literacy for Business Administration

Stefan Dreisiebner

University of Graz, Austria
stefan.dreisiebner@uni-graz.at

Abstract

Information Literacy (IL) is “the set of integrated abilities encompassing the reflective discovery of information, the understanding of how information is produced and valued, and the use of information in creating new knowledge and participating ethically in communities of learning” (American Library Association, 2016). IL as a social key competence is particularly essential in post-secondary education and research. According to many studies (e.g. Katz, 2007; Rubinić et al., 2013), student’s information literacy levels are generally low. IL also includes subject-specific skills; e.g. relevant resources are different among disciplines (Johnston & Webber, 2003; Bundy, 2004). Analyses show the need for an attractive tutorial which can be developed on a broad basis of available material. Accordingly, an extension to a Massive Open Online Course (MOOC) for developing information literacy focusing on higher education students shall be developed and evaluated, that focuses on information skills needed in Business Administration. MOOCs are online courses with no entry barriers open for unlimited participation (van Treek et al., 2013).

Keywords: information literacy; MOOC; subject-specific skills; business administration

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 298–302.

1 Methodology

The thesis will follow the design science approach, which seeks to extend the boundaries of human and organizational capabilities by creating new and innovative artefacts (Hevner et al., 2004; Becker et al., 2009; Mertens, 2010; Heinrich, 2011). Accordingly, a generic artefact will be created by developing a subject-specific MOOC extension. The phases as defined by Becker et al. (2009) are analysis, design, evaluation and diffusion.

The main step of the analysis phase will be research on existing online learning material on subject-specific IL contents for Business Administration. For this purpose, an online research will be done on all available material in German that is online, with a special focus on MOOCs. Additionally, discipline-specific standards (e.g. Cunningham, 2003) and books (e.g. Herbig, 2011) shall be analyzed.

Out of the analysed material a collection of subject-specific skill-sets and material shall be created. This material will be verified and probably further extended by semi-structured expert interviews (Liebold & Trinczek, 2009; Meuser & Nagel, 2009). The interviews should involve specialist staff on the discipline from libraries, professors teaching IL for Business Administration and professors that essentially need IL skills in their teaching within Business Administration (e.g. professors teaching master seminars). The resulting material will build the basis of the subject-specific MOOC extension. A special emphasis will be put on the didactic concept. This will follow previous experience with MOOC development (Lackner et al., 2014) and several state-of-the-art approaches like the First Principles of Instruction (Merrill, 2002) and the revised Bloom's taxonomy (Krathwohl, 2002).

The content will be published within an existing MOOC as described in the Expected Outcome and evaluated while applied in practice. Therefore, several approaches will be used. First of all, a thinking aloud analysis (van Someren et al., 1994) will be applied to test the usability and clearness of the solution. In the test the persons will solve several tasks in the MOOC while saying their thoughts. This, along with the software interactions and a short follow-up interview will be evaluated. Second, the gains in IL of the participants will be measured. This will be done by using a questionnaire testing several real-world problems related to the discipline that will be handed out before and after attending the MOOC. Third, a questionnaire will be included in the subject-specific extension asking the users about their satisfaction.

The diffusion of the results (Becker et al., 2009) is guaranteed by publishing a publicly available online learning resource, along with several dissemination activities.

2 Expected outcome

The resulting MOOC extension will be part of a MOOC that is developed within the EU project Information Literacy Online (ILO), which started in November 2016. The ILO project will at first concentrate on information literacy elements which are relevant for all subjects/disciplines. A central innovative approach of this MOOC will be the implementation of technology based assessment components which allow students to get feedback on their learning success. The MOOC content will be developed in six European cultural and language groups: English, German, Spanish, Catalan, Slovenian and Croatian. The multilingual approach will not only consider formal translation but also cultural-specific differences in the various realizations.

Within the project also subject-specific extensions for Business Administration and Psychology shall be developed. The responsibility for the extension on Business Administration is under the sole responsibility of this author. The results will show how a subject-specific extension of a MOOC could look like and how this could be integrated in “generic” teaching material on IL. It will build the foundations for more effective teaching offers in IL in the future, which can be integrated more flexible into existing teaching offers due to the modular design with subject-specific extensions. Additionally, it brings value for the society by publishing the results as Open Educational Resource.

Acknowledgement

The author would like to thank Christian Schlögl (University of Graz) for supervising this thesis.

References

- American Library Association, ACRL Board (2016): Framework for information literacy for higher education. Chicago, ALA. <http://www.ala.org/acrl/standards/ilframework> <30.11.2016>
- Becker, Jörg, Helmut Krcmar, and Björn Niehaves (2009): *Wissenschaftstheorie und gestaltungsorientierte Wirtschaftsinformatik*. Heidelberg: Physica-Verlag. [doi:10.1007/978-3-7908-2336-3](https://doi.org/10.1007/978-3-7908-2336-3)
- Bundy, Alan (2004): *Australian and New Zealand Information Literacy Framework: Principles, Standards and Practice* (2nd ed.). Adelaide: Australian and New Zealand Institute for Information Literacy.
- Cunningham, Nancy A. (2003) Information Competency Skills for Business Students. In: *BRASS Business Reference in Academic Libraries Committee* 1 (1): 1–4. <http://www.ala.org/rusa/sections/brass/brasspubs/academicbrass/acadarchives/volume1numnum1/academicbrassv1> <30.11.2016>
- Heinrich, Lutz J. (2011): *Geschichte der Wirtschaftsinformatik*. Berlin, Heidelberg: Springer. [doi:10.1007/978-3-642-16859-8](https://doi.org/10.1007/978-3-642-16859-8)
- Herbig, Albert F. (2011): *Informationskompetenz Wirtschaft: Erfolgreiche Informationsrecherche für das betriebswirtschaftliche Bachelor- und Masterstudium*. Norderstedt: Books on Demand.
- Hevner, Alan R., Salvatore T. March, Jinsoo Park, and Sudha Ram (2004): Design Science in Information Systems Research, In: *MIS Quarterly*, 28 (1): 75–105. [doi:10.2307/25148625](https://doi.org/10.2307/25148625)
- Johnston, Bill, and Sheila Webber (2003): Information Literacy in Higher Education: A Review and Case Study. In: *Studies in Higher Education* 28 (3): 335–352. [doi:10.1080/03075070310000113441](https://doi.org/10.1080/03075070310000113441)
- Katz, R. Irvin (2007): Testing information literacy in digital environments: ETS's iSkills assessment. In: *Information Technology and Libraries*, 26 (3), 3–12. <http://ejournals.bc.edu/ojs/index.php/ital/article/viewFile/3271/2884> <30.11.2016>
- Krathwohl, David R. (2002): A Revision of Bloom's Taxonomy: An Overview. In: *Theory Into Practice* 41 (4): 212–218.
- Lackner, Elke, Michael Kopp, and Martin Ebner (2014): How to MOOC? – A Pedagogical Guideline for Practitioners. In: Ion Roceanu (Ed.): *Proceedings of the 10th International Scientific Conference "eLearning and Software for Education", Bucharest, April 24–25, 2014*. Bucharest: Editura Universitatii Nationale de Aparare "Carol I".

- Liebold, Renate, and Rainer Trinczek (2009): Experteninterview. In: Stefan Kühl, Petra Strodtholz, and Andreas Taffertshofer (Eds.): *Handbuch Methoden der Organisationsforschung* (pp. 32–56). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Merrill, M. David (2002): First Principles of Instruction. In: *Educational Technology Research and Development* 50 (3): 43–59. [doi:10.1007/BF02505024](https://doi.org/10.1007/BF02505024)
- Mertens, Peter (2010): Anspruchsgruppen der gestaltungsorientierten Wirtschaftsinformatik. In: Hubert Österle, Robert Winter, and Walter Brenner (Eds.): *Gestaltungsorientierte Wirtschaftsinformatik: Ein Plädoyer für Rigor und Relevanz* (pp. 19–25). Nürnberg: infowerk.
- Meuser, Michael, and Ulrike Nagel (2009): Das Experteninterview – Konzeptionelle Grundlagen und methodische Anlage. In: Susanne Pickel, Gert Pickel, Hans-Joachim Lauth, and Detlef Jahn (Eds.): *Methoden der vergleichenden Politik- und Sozialwissenschaft* (pp. 465–479). Wiesbaden: VS Verlag für Sozialwissenschaften. [doi:10.1007/978-3-531-91826-6_23](https://doi.org/10.1007/978-3-531-91826-6_23)
- Rubinić, Dora, Stričević Ivanka, and Juric Mate (2013): Information Literacy Course – The Perception of Students and Professors: University of Zadar Case. In: Kurbanoglu Serap et al. (Eds.): *Worldwide Commonalities and Challenges in Information Literacy Research and Practice: European Conference on Information Literacy, ECIL 2013, Istanbul, Turkey, October 22–25, 2013, Revised Selected Papers* (pp. 528–534). Cham et al.: Springer International Publishing. <http://goo.gl/oa4F1C> <30.11.2016>
- van Someren, Maarten W., Yvonne F. Barnard, and Jacobijn A. C. Sandberg (1994): The Think Aloud Method: A Practical Guide to Modelling Cognitive Processes. Department of Social Science Informatics, University of Amsterdam. [doi:10.1016/0306-4573\(95\)90031-4](https://doi.org/10.1016/0306-4573(95)90031-4)
- van Treeck, Timo, Klaus Himpsl-Gutermann, and Jochen Robes (2013): Offene und partizipative Lernkonzepte. E-Portfolios, MOOCs und Flipped Classrooms. In: Martin Ebner and Sandra Schön (Eds.): *Lehrbuch für Lernen und Lehren mit Technologien* (2nd ed.) (pp. 287–300). Berlin: epubli.

Mobile Privacy and Apps: Investigating User Behavior and Attitude

*Stefanie Havelka*¹

Humboldt-Universität zu Berlin
Institut für Bibliotheks- und Informationswissenschaft
Unter den Linden 6, D-10099 Berlin, Germany
shavelka@gmail.com

Abstract

My dissertation investigates privacy as it relates to user behavior and attitude on mobile devices. I will examine how users' understanding of mobile privacy differs from culture to culture. More specifically, I will use ethnographic methodologies to compare US and German library/information science students.

Keywords: mobile privacy; privacy; apps; mobile devices; user behavior; attitude; United States; Germany; library and information science students; ethnography

1 Introduction

Using applications (“apps”) on mobile devices (i.e., smartphones, phablets and tablets) to access information has increasingly become the norm in our global society, replacing desktops and laptop computers as users' preferred access to the internet. Apps can automatically capture a broad range of user

¹ Assistant Professor | Web and Mobile Service Librarian | Lehman College | The City University of New York | Bronx, New York | USA

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 303–307.

information that is already stored on the device, including the user's precise location, phone number, list of contacts, call logs, unique device identifiers, and more. Because today's mobile devices have developed such wide-ranging capabilities in a short period of time, preserving user privacy has become a "hot button" issue. But many users (myself included) are often unaware of or indifferent to what types of personal data certain apps can access and what is shared with the device manufacturer, app vendor, and even third parties. "This is typically seen as an expression of the 'privacy paradox' where intentions and behaviors around information disclosure often radically differ" (Shklovski et al., 2014: 2347).

2 Preliminary literature review

A preliminary literature review assessing 88 articles from 2010 onward found a wealth of articles addressing social networking and user privacy behavior – especially from the perspective of computer science and/or human-computer-interaction disciplines. But when it comes to privacy and mobile devices, the majority of research has been done by law scholars and computer scientists, as well as government/policy stakeholders and businesses, and user behavior as it relates to mobile privacy is, so far, underrepresented. The same can be said for the information and library science field: only two articles address privacy. The first, by Magnuson (2011) posited that online regulation and privacy management are an integral part of information literacy skills in academic libraries. The second, by Cyrus and Baggett (2012), explored the relationship between libraries and privacy, the emergence of mobile technologies, and how librarianship is challenged by user privacy issues. One of the more recent articles does not look at mobile privacy per se, but it is the only research that juxtaposes mobile-information-seeking behavior across different cultures (Lee & Song, 2015).

3 Research question

Are there differences in the mobile privacy user behaviors and attitudes of American and German library and information science students?

In order to establish the correct context for my research, I need to clarify the following:

- (a) Mobile privacy entails personal data and information being accessible or transferred on mobile devices to device manufacturers, app developers and/or other third parties. It does not include data privacy, online privacy on the desktop or laptop.
- (b) User behavior is defined as how a study participant utilizes mobile devices, apps, and websites in their everyday life.
- (c) Attitude investigates the feelings and established perceptions of study participants.

4 Methodology

My dissertation will use ethnography as its qualitative method. “Ethnographers typically describe a particular situation or process by asking multiple people about it, and by analyzing multiple types of data, such as interviews, direct observation, photographs, journals, or cultural artifacts” (Duke & Asher, 2012: 3). For Clifford Geertz (1973), Ethnographic description is interpretive of the flow of social discourse and interpreting it involves preserving the “said” of such discourse. For Seadle (2011), library-based ethnography is comprised of three components: a valid research question; data gathering to answer it; and conducting the analysis of gathered data to answer the research question persuasively.

5 Data collection

Data will be collected via interviews which entails also an experiment with participant observation. Fieldwork will be conducted in Germany and the United States. Participants for this study will be:

- **In Germany:** Ten students currently enrolled in *Masterstudiengang Bibliotheks- und Informationswissenschaft* or *Fernstudium* (Master in Library and Information Science or Distance Study) at the *Institut für Bibliotheks- und Informationswissenschaft (Berlin School of Library and Information Science)*, *Humboldt-Universität zu Berlin (Humboldt University in Berlin)*.
- **In the United States:** Ten students currently enrolled in the *Master of Information Program* concentrating in Library and Information Science at *Rutgers University*.

Prior to the beginning of the study, all participating students will be asked to fill out a consent form. Furthermore, students will be informed about the confidentiality and anonymity of the collected data. Information that reveals the identity of the respondent will be omitted from research results.

Interviews will be video recorded. All collected data files will be transcribed and stored securely. A software program such as ATLAS.ti (<http://atlasti.com/>) will be used to facilitate data analysis.

References

- Cyrus, J. W. & Baggett, M. P. (2012): Mobile technology: Implications for privacy and librarianship. In: *The Reference Librarian*, 53 (3), 284–296.
- Duke, L. M. & Asher, A. D. (2012): *College libraries and student culture: What we now know*. Chicago: American Library Association.
- Geertz, C. (1973): *The interpretation of cultures: Selected essays*. New York: Basic Books.
- Lee, J. & Song, Y. (2015): Mobile information-seeking behavior: A comparative study. In: *IFLA Journal*, 41 (2), 153–161.
- Magnuson, L. (2011): Promoting privacy Online and reputation management as an information literacy skill. In: *College & Research Libraries News*, 72 (3), 137–140.

Seadle, M. (2011): Research rules for library ethnography. In: *Library Hi Tech*, 29 (3), 409–411.

Shklovski, I., Mainwaring, S. D., Skúladóttir, H. H. & Borgthorsson, H. (2014): Leakiness and creepiness in app space: Perceptions of privacy and mobile app use. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (pp. 2347–2356). ACM.

Visualization for Text Mining in the Digital Humanities

Empowering Researchers to Use Advanced Tools for Text Mining

Julian Hocker

DIPF, Germany
Julian.Hocker@dipf.de

Abstract

In this PhD thesis, a visual interface for text analysis and text mining in the digital humanities (DH) will be developed. Text analysis is a crucial task in the DH, but advanced text mining technologies like topic modeling or clustering are difficult to use for most researchers. My work bridges this gap using visualizations. To ensure an adequate usability of visualizations for epistemological practices, the visualizations will be realized with researchers in an agile and participatory approach.

Keywords: visualization; text mining; usability; digital humanities

1 Introduction

The potentials of DH are often seen in the analysis of large corpora of texts. On the one hand there are very advanced tools for clustering and topic modelling available, but on the other hand most of the researchers in Humanities are lacking the capacities to use text mining or machine learning. Wagstaff (2012) already addresses this issue concerning machine learning in general.

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 308–313.

How can quantitative *distant reading* (Moretti, 2016) tools become adequate epistemological tools (Ramsay & Rockwell, 2012) for humanities research in practice? How can these methods be combined with qualitative analysis? In my dissertation I follow these questions by using visualizations as a possibility to create an epistemological interface. Visualizations are regarded as a means for supporting intuitiveness and usability, in order to help people with less knowledge to be able to use these techniques. This needs to consider the concrete context in use, in order to design the adequate, new capacities. In my case, visualizations are used to enhance epistemological practices in Humanities.¹

A research project² in historical educational research will be used as a case study for designing and evaluating a visual interface. The semantic research environment *Semantic CorA* (Schindler et al., 2012), which is based on *Semantic Media Wiki* and already offers tools for qualitative research, will be used as the central platform. There are only few approaches to add text analysis functions to Wikis (Witte & Sateli, 2014; Mehler et al., 2016), but all are lacking explorative visualizations and a combination with tools from close and distant reading.

My research questions are: How can a visualization interface enhance the epistemological practices in humanities? How can methods of close and distant reading be thoroughly combined? Which research capacities need to be addressed and designed to enhance humanities research? How can people with little or no knowledge about text mining be encouraged to use these methods? How can we address and create capacities for the usage of these tools in these humanities research groups?

2 State of the art

Jänicke et al. (2016) have conducted a survey of published papers for visualization in the DH. They see an emerging use of combining close and distant

¹ Visualization is a very vivid field, there is some work done in DH, for example a new workshop format starting in October 2016 at the IEEE VIS conference (<http://vis4dh.com/>).

² The project is called “Abiturprüfungspraxis und Abituraufsatz 1882 bis 1972” together with partners from Humboldt University Berlin and the KIT.

reading methods and introduce a taxonomy, where visualizations can be grouped based on the task they fulfil instead of the classical grouping in close or distant reading.

Researchers in DH are often not familiar with text mining technology and therefore not aware of what is possible and what is not. Because of this, visualizations should be developed together with the researchers. The developer also has to keep in mind that a visualization is not the end of the research, but is used to generate new research questions and lets the researcher dive deeper into the analysis (Jänicke, 2016). This is similar to the basic idea of explorative visualizations. Ramsay (2007) states also that the text analysis should be seen as a tool which allows analyzing the material.

There are different tools that offer text analysis methodologies for researchers in DH. But most of these tools either focus on close reading approaches (cf. Cheema et al., 2016) or do not enable users to use advanced text mining methods like clustering or topic modelling (Rockwell & Sinclair, 2016).³ Some tools are also bound to certain corpora and therefore not reusable for other projects or corpora.

For creating visualizations with a focus on epistemological practices, the classical approach in context of usability is the visual information seeking mantra (Shneiderman, 1996). Sedlmair (2012) proposes a metric for user-centered creation of visualizations. It has three main stages, *precondition*, which is about the basic goals of the project, goals and set up, the *core stage*, which involves discovery what people want, the design and iterations of implementing code and deploying it in order to get feedback. The last stage is the *analysis* where the researchers should reflect the coding and publish the results at the end of the project. This categorization makes a lot sense because it brings agile software development to the creation of visualizations. It also means giving researches the ability to create visualizations and come to their own analysis, so researchers in the DH can be able to have a “humanistically informed theory of the making of technology”, as Drucker (2012) demands. These tools have to be developed together with users from the DH in order to be useful (Borgman, 2009).

Kath et al. (2015) claim that the visualization of data is already an interpretation of data. Therefore it is necessary to make clear how researchers select the data, pre-process it using text mining algorithms and also how the visualizations are done. Some tools like *iPython* (Perez & Granger, 2007)

³ Voyant Tools, available at <http://www.voyant-tools.org>

and R^4 provide this and can be also used with little knowledge about programming, a similar tool especially for research in DH is *Voyant Notebook* (Rockwell & Sinclair, 2013). All the tools also make obvious which commands have been applied to text, but these tools have the disadvantage that users need basic coding skills and this is often not the case.

3 Research design and methods

In order to make the visualizations of an epistemological tool suitable for researchers in DH, a mix method approach will be realized in three steps. In a first step, Sedlmair's requirement analysis approach will be followed for adjusting techniques of natural language processing (NLP) to the needs of humanities researchers and for identifying the concrete problems by using these tools. Therefore there will be a participant observation and a tool analysis. Based on these results, there will be expert interviews.

The development will focus on a participative and agile approach with a close connection to researchers in digital humanities. Several visits with close contact to researchers and evaluation of prototypes are planned. In a last step the developed tool will be evaluated by an expert test, which addresses the concrete new research capacities like exploration and interaction with the data, with researchers in the field of the history of education or DH.

4 Expected results

I expect my results to show an improvement for the exploration of massive text data for the non-experts through visualizations as an epistemological tool. This contains the act of interpretation of data enrichment as well as the choosing of data units to analyze. Combined with the research environment *Semantic CorA* I introduce a tool for qualitative and quantitative analysis of text corpora.

4 <https://www.r-project.org>

Explorative visualizations, as proposed by Jürgens et al. (2015) for patent retrieval might be a good starting point and there might be also a need to educate the DH researchers in text mining methods.

References

- Borgman, C. L. (2009): The digital future is now: A call to action for the humanities. In: *Digital humanities quarterly*, 3 (4).
- Cheema, M. F., Jänicke, S. Scheuermann, G. (2016): AnnotateVis: Combining Traditional Close Reading with Visual Text Analysis. In: *Workshop on Visualization for the Digital Humanities, VisWeek, 2016*.
- Drucker, J. (2012): Humanistic theory and digital scholarship. In: Matthew K. Gold (Ed.): *Debates in the digital humanities* (pp. 85–95). Minneapolis, London: University of Minnesota Press.
- Kath, R., Schaal, G. S. & Dumm, S. (2015): New Visual Hermeneutics. In: *Zeitschrift für germanistische Linguistik*, 43 (1), 27–51.
- Jänicke, S. (2016): Valuable Research for Visualization and Digital Humanities: A Balancing Act. In: *Workshop on Visualization for the Digital Humanities, VisWeek*.
- Jänicke, S., Franzini, G., Cheema, M. F. & Scheuermann, G. (2016): Visual Text Analysis in Digital Humanities. In: *Computer Graphics Forum*. [doi:10.1111/cgf.12873](https://doi.org/10.1111/cgf.12873)
- Jürgens, J. J., Mandl, T., Womser-Hacker, C. (2015): Visualizing Query Comparisons in Patent Retrieval Systems. In: *Proceedings of the Second International Workshop on Patent Mining and Its Applications (IPaMin 2015)*. http://ceur-ws.org/Vol-1437/ipamin2015_paper5.pdf <27.10.2016>
- Mehler, A., Gleim, R., Hemati, W., Uslu, T. & Eger, S. (2016): Wikidition: Automatic lexiconization and linkification of text corpora. In: *it – Information Technology*, 58 (2), 70–79.
- Moretti, F. (2013): *Distant Reading*. London, New York: Verso.
- Pérez, F. & Granger, B. E. (2007): IPython: a system for interactive scientific computing. In: *Computing in Science & Engineering*, 9 (3), 21–29.
- Ramsay, Stephen (2007): Algorithmic criticism. In: Ray Siemens and Susan Schreiban (Eds.): *A Companion to Digital Literary Studies*. Malden, Mass.: Blackwell. <http://digitalhumanities.org/companion/view?docId=blackwell/9781405148641/9781405148641.xml&doc.view=print&chunk.id=ss1-6-7&toc.depth=1&toc.id=0> <24.10.2016>

- Ramsay, S. & Rockwell, G. (2012): Developing Things: Notes toward an Epistemology of Building in the Digital Humanities. In: Gold, Matthew (Ed.): *Debates in the Digital Humanities* (pp. 75–84). Minneapolis, MN: University of Minnesota Press.
- Rockwell, G. & Sinclair, S. (2016): *Hermeneutica: Computer-Assisted Interpretation in the Humanities*. Cambridge, Mass.; London: The MIT Press.
- Schindler, C., B. Ell, M. Rittberger (2012): Intra-linking the Research Corpus. Using Semantic MediaWiki as a lightweight Virtual Research Environment. In: Meister, Jan Christoph et al. (Eds.): *Digital Humanities 2012* (pp. 359–362). Hamburg: Hamburg University Press.
- Sedlmair, M., Meyer, M. & Munzner, T. (2012): Design Study Methodology: Reflections from the Trenches and the Stacks. In: *IEEE transactions on visualization and computer graphics*, 18 (12), 2431–2440.
- Shneiderman, B. (1996): The eyes have it: a task by data type taxonomy for information visualizations. In: *1996 IEEE Symposium on Visual Languages* (pp. 336–343).
- Sinclair, Stéfan and Geoffrey. Rockwell (2013): Voyant Notebooks: Literate Programming, Programming Literacy. In: *Journal of Digital Humanities*, 2 (3).
- Wagstaff, Kiri (2012): Machine Learning that Matters. In: *Proceedings of the 29th International Conference on Machine Learning, Edinburgh, Scotland, UK, 2012*. <http://icml.cc/2012/papers/298.pdf>
- Witte, R. & Sateli, B. (2014): Adding Natural Language Processing Support to your (Semantic) MediaWiki. In: *The 9th Semantic MediaWiki Conference (SMWCon Spring 2014)*.

Development of Knowledge Management Performance Metrics for Enterprise Social Software

Fabian Odoni

Swiss Institute for Information Research (HTW Chur), Switzerland
fabian.odoni@htwchur.ch

Abstract

This is a summary of a planned dissertation research. Goal is to develop knowledge management metrics that can be used to measure the performance of enterprise social software. To achieve this goal an ontology will be developed showing all entities and properties of enterprise social software. Indicators will then be extracted and new ones generated using natural language processing, graph theory and psychometrics to enrich the data. In the end a visualisation of the indicators (knowledge map) will be developed to present and verify the indicators.

Keywords: information retrieval; data visualisation; knowledge management; software development, metrics; ontology; natural language processing; graph theory; social enterprise software

1 Introduction

Enterprise social software offer methods that are designed to increase efficiency and effectiveness in organisations by allowing employees to share and search for knowledge (e.g., Microsoft Yammer, IBM Connection, Novell Vibe ...). They are often designed to emulate Facebook by providing a

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 314–320.

content stream that can be commented and by offering the users the possibility to create groups of interest (often referred to as spaces) in which they can discuss or work on different topics. Additionally they often offer wikis, document and/or file management solutions, Questions and Answers modules and other functionalities. Although these platforms claim to promote efficiency and effectiveness they do not yet offer any method that allows an organisation to verify these claims or detect other non-technical problems that prevent effective knowledge sharing (e.g., cultural problems).

2 Research interest

Inspired by projects such as stackoverflow.com, that offer a wide range of metrics to allow the users to inform themselves on their current state of knowledge and compare themselves with other users, the aim of this dissertation research is to identify indicators that allow the quantification and presentation of the state of knowledge management of platforms using enterprise social software.

Following questions will guide the research:

1. Who seems to know what based on their actions using enterprise social software?
2. Who is sharing what knowledge with whom using enterprise social software?
3. What kind of key players can be found on enterprise social software platforms and what do they do?
4. Why is who sharing or not sharing their knowledge using enterprise social software?

3 Research design

To answer these questions following steps are planned.

3.1 Literature research

A thorough literature research will provide the foundation of this dissertation. The focus will lie on ontologies, enterprise social software used as knowledge management tools, natural language processing, graph analysis and psychometrics.

3.2 Identification and development of enterprise social software indicators

Following steps will be made in order to identify and develop enterprise social software indicators that are able to represent the state of knowledge management in such an environment.

3.2.1 Analysis of enterprise social software

The goal of this step is to gather knowledge on different enterprise social software solutions. In order to do that, a number of different products must be identified. These are then dissected to find all the different tools and methods they offer (e.g., posting-streams, wikis, spaces ...). These components will then be grouped based on their function as a knowledge management tool and described. The result will be a comprehensive table listing all possible knowledge management methods and their function used in enterprise social software.

3.2.2 Creating an ontology to represent knowledge management with enterprise social software solutions

After all the methods and functions are known, an ontology will be created that can be used as a map to navigate the possibilities of enterprise social software. This will allow the discovery of relevant indicators and support the development of new indicators.

To create such an ontology, first all the entities and their properties must be identified. These will include the users of the system as well as the different objects that are created and presented with enterprise social software installations. Were possible, entities will be hierarchically ordered to identify

common properties and functions shared between them. After that, all predicates will be identified and assigned.

Though an ontology will be able to show all possibilities that are available using enterprise social software further research must be done to identify if an ontology can represent the implications of different possible values that entity properties may have.

Thus, at this point, it is unclear if an ontology is sufficient to represent the relations between the entities of an enterprise social software product or if at this stage, a different representation e.g., a class diagram as known from software development, must be used in order to represent all the possibilities available. This might be the case if for example methods and conditions are discovered that dynamically change the relationships between the entities depending on their values that cannot be described using an ontology.

3.2.3 *Defining and developing relevant indicators*

Based on the previous work, indicators can be defined that allow the quantification of certain states of knowledge management using enterprise social software (how much knowledge is shared by whom, who knows what, what is missing ...). These indicators will also be defined following the basic system of bibliometry e.g., by Havemann (2009).

Additionally new indicators will be developed by enriching the currently available data. At this point following methods are considered in order to achieve this goal:

1. *Natural language processing*: Using natural language processing methods keyword extraction and topic identification can be utilised to generate data that can be used as properties for entities or relationships.
2. *Graph Theory*: Using graph theory new relationships and indicators can be identified. By building a graph based on the ontology (or another suitable representation thereof e.g. class diagram) graph theory methods can be applied to find additional data that can be used to generate new indicators e.g., the centrality and importance of an actor as described by Dehmer et al. (2015) and Cross and Parker (2004).
3. *Psychometrics*: Using psychometrics the intention or personality of the users can be approximated. This data can be used to generate additional properties that might allow an explanation behaviour or situations.

It is not the goal to develop entirely new methods and algorithms to enrich the data but to find, evaluate and use existing methods.

The results of these additional data enrichment steps must be evaluated and if deemed to be adequate will be implemented into the representation methods already developed at this point (ontology, class diagram ...).

The result of this step will produce a set of indicators (numbers, relationships ...) and describe their value as well as how they can be computed and what is needed to do so. Grouping and ordering of the indicators and measures e.g., by subject and content criteria (Hummel, 2003: 555), will be applied where and if possible.

Additionally a framework will be programmed that can extract and compute the indicators. To do so, the right technologies must be identified and mastered so that they can be implemented into the framework.

3.3 Development of knowledge maps to navigate and explore the indicators and relationships

In order to utilise these indicators and relationships a visual representation will be developed by creating knowledge maps displaying specific aspects of the data collected. This will allow a first verification of the usefulness of the indicators as well as provide an easy summary of the data that can then be implemented as knowledge management tool into enterprise social software.

Other forms of visual presentation of data will be considered and partially developed but the focus will lie on presentation forms that allow the visualisation of data of multiple entities and the relationships among them.

4 First results

Although this dissertation is just at the beginning, first steps were made within the context of an ongoing research project. For this project, a co-author analysis of wiki pages was computed and the relationship between the authors was visualised (cf. fig. 1).

These first results were able to show some interesting relationships between the authors and already at this stage, it was a big success among the project partners.

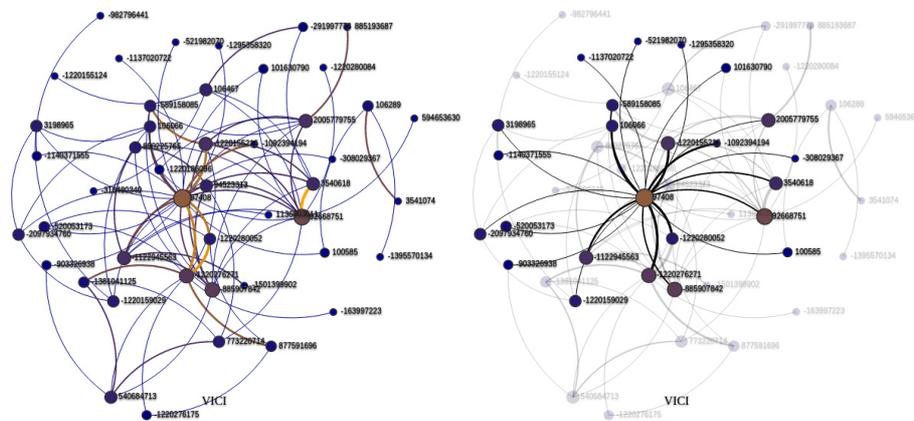


Fig. 8 VICI co-author visualisation of a Confluence database dump and highlighted nodes and links on mouse-over (displaying anonymized user names)

5 Conclusion

This dissertation research is only just beginning and thus many points are still unclear. By embedding the research in ongoing projects a number of aspects and possibilities are still subject to change in order to adhere to the project's goals. This presentation should serve as a first opportunity to gather feedback from researchers and students not involved in the project and thus generate new ideas and point out possible problems from an outside point of view.

References

- Cross, Rob; Parker, Andrew (2004): *The hidden power of social networks. Understanding how work really gets done in organizations* (2nd printing ed.). Boston, Mass.: Harvard Business School Press.
- Dehmer, Matthias; Emmert-Streib, Frank; Pickl, Stefan (2015): *Computational Network Theory: Theoretical Foundations and Applications*. Weinheim: Wiley-VCH.

Havemann, Frank (2016): *Einführung in die Bibliometrie*. Berlin: Gesellschaft für Wissenschaftsforschung.

Hummel, Thomas (2003): Quellen und Elemente von Informationssystemen des Controlling. In: Steinle, Claus; Bruch, Heike (Eds.): *Controlling – Kompendium für Ausbildung und Praxis*. Stuttgart: Schäffer-Poeschel.

Information Visualization of Environment Maps for Complex, Relational Data for Better Reflected Decision Making on the Example of Food

Veit Stephan

University of Regensburg, Germany
veit.stephan@ukr.de

Abstract

My dissertation aims at developing an application that presents all the information a user needs to find orientation in his environment. All surrounding places and their characteristics shall be perceived and compared in a fast way. The goal is to implement an overview map application to reduce information complexity of one's spatial location by using appropriate interaction and visualization methods.

Keywords: user interfaces; visualization, food environment maps

1 Introduction

Due to a growing amount of data and the problems related to this trend (Lange, 2009), there is a need for user centered information visualization software (Macaulay et al., 2009), optimized for target group needs – software that enables a better handling of large amount of information (Gantz, 2008). One big challenge by working with webGIS, in regard to location based applications (LBAs), is the complexity of data (Longley, 2011) and to present

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 321–328.

them in an overview in order to enable users to rapidly and cognitively process the information (Xu et al., 2011).

1.1 Food environment as an example

Health research shows that for individual nutrition behavior environment and circumstances act as influential forces (Loss & Leitzmann, 2011). This aspect leads to the idea of gathering information about explicit indicators of a food related point of interest (PoI).

1.2 Visualization of complex data for food environment maps

A food environment application with multi parametrical data has to be based on an interactive information system reliably presenting the density of information, through fast and simple navigation (cf. Mazza, 2009: 105–124), overview, filtering and detail information views (Shneiderman, 2003).

1.3 Objectives

The task is to develop an application which enables users to more easily classify and recognize their surroundings according to personal assumptions. Therefore the user shall be able to perceive all relevant information about surrounding PoIs and its nutrition related aspects fast and intuitively. The aim is to find out how concepts of interaction and visualization for overview map applications can reduce the complexity of the information contained in one's surrounding.

2 Materials and Methods

To examine how to prepare data best for interactive environment maps applications, during the concept stage elaborating descriptive/explorative pre-studies will be executed. These results, in addition to a short review of the further developments of the reference model of information visualization, shall lead to the development of a new model. By agile software development in dependence on the DIN EN ISO 9241-210 (2010) I create an alpha

release of the food environment map. The following step is to measure an improved beta release by asynchronous remote usability test in combination with A/B tests.

3 Previous Results

3.1 Model formation and und concept

The analysis of the achievements of the reference model of information visualization shows by starting from Salton and McGill and their concept of “iterations during information retrieval” (cf. 1983: 237) looking from Ellis, Cox and Hall describing the “process model” (cf. 1993: 356–356) to for example Belkin’s “model of episodes” (cf. 1996: 25–31) along to the “cognitively extended model for information retrieval” of Landwich from 2007 (fig. 1) that the user itself gains an increasingly prominent role during the modeling. Thus, one arrives at following hypothesis: By modeling information retrieval, there is a tendency to focus more closely on the involvement of the user and its cognitive ability to process information as a part of the retrieval procedure.

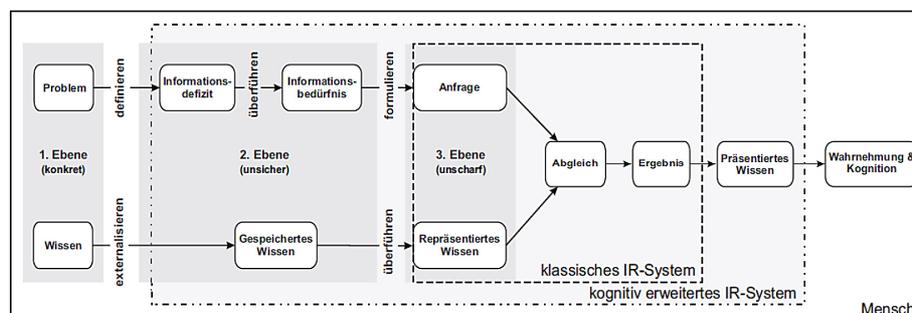


Fig. 1 Cognitively extended model for information retrieval (Landwich, Hemmje & Fuhr, 2007: 328)

Based on this hypothesis, a model has been created that maximizes the match between the user’s needs and the characteristics of one unit of retrieval (UoR). The intention is to create an application which makes it possible to present multi parametrical, relational data in a specific context by achieving a

fast and comprehensible success in searching, assessing and preferencing the data. A major thought at this juncture is to focus the role of the user and its ability to interpret and intuitionally perceive the information and therefore influence the selection and emphasis of it (fig. 2).

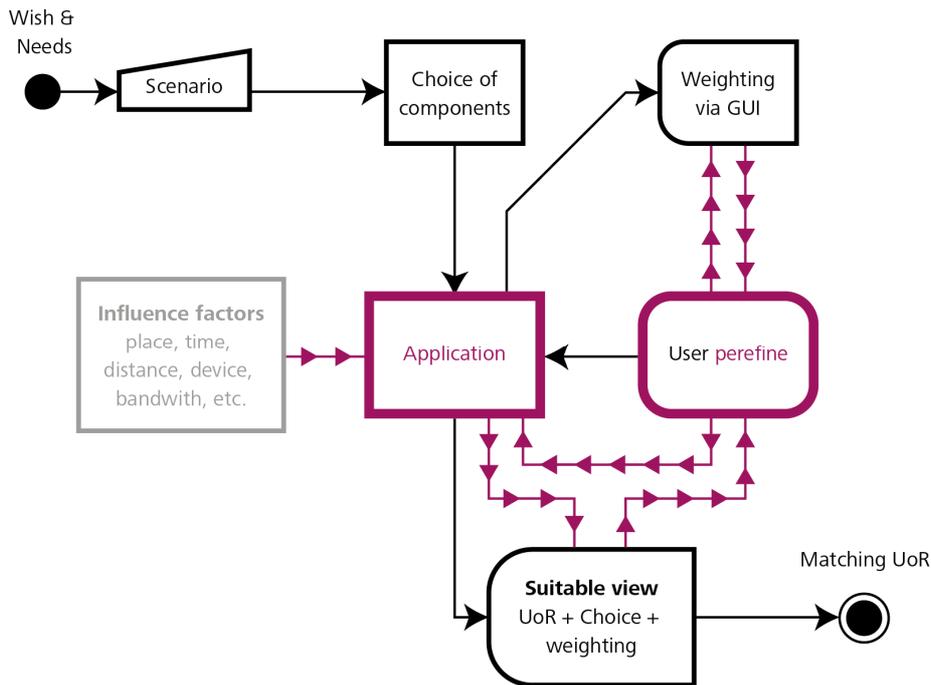


Fig. 2 Model of processes

This continuous user interaction is called “perfine”. As an acronym for the coalition of the processes of [cognitive user] per[ception], ref[lection] [and] in[t]e[r]action].

As a consequence, the idea of an interaction and visualization model evolves, that iterates from a user centered perspective over following steps:

1. A user need is abstracted as
2. A scenario, which leads to a
3. Selection of criteria and components.
4. That will be assessed.
5. Out of which a visual structure will be created and
6. Presented in a view.
7. All the results will be reflected.

These steps are put in order in a rotation loop to demonstrate how closely all these elements are connected to the user's cognitive perception and ability to reflect.

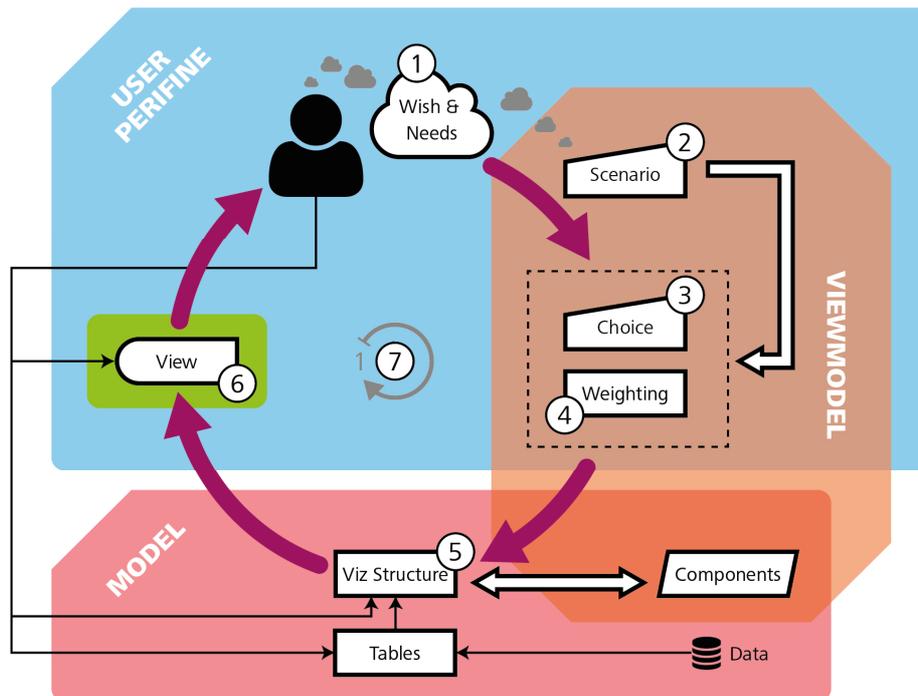


Fig. 3 Concept model of interaction and visualization – as projection on MVVM (cf. Kühnel, 2013: 1091–1096)

3.2 Prototypes

A prototype like alpha implementation (Stephan, 2016b) with crawled data from the Google maps API (Google Developers, 2016) combined with randomized data per UoR provided different visualization methods in order to prove the feasibility of each. These examinations pointed out that problems can occur with the dataset which lead to bias. Incidentally, I tried to improve the application adding different measuring scales and replacing the data by parts of the “Yelp Dataset Challenge” dataset (Yelp, 2016). In this way it was possible to create more mock-ups (Stephan, 2016a) and simultaneously, the beta release development has been pushed.

4 Research challenges and outlook

Finally I will evaluate the beta release prototype by asynchronous remote usability tests. The plan is to compare the application in an A/B test against a standard tag-based application and measure the usability by using the meCUE (Minge & Thüring, 2009) questionnaire. It is to clarify here which requirements the tag based applications needs to fulfil to minimize bias. Furthermore, the recruitment of a sufficiently large test group poses a challenge. Generally the requirement to guarantee the identification of major key facts for evaluation has to comply.

In consideration of good usability of the model more use cases shall be illustrated. Through transforming it into other context areas like i.e. noise and silence, working conditions or education the adaptability to meet the different needs shall be proved. It would be the aim to assist the reflections and comparisons of users with utilizable systems for complex decision criteria.

References

- Belkin, Nicholas J. (1996): Intelligent information retrieval: whose intelligence? In: Krause, Jürgen; Herfurth, Matthias; Marx, Jutta (Eds.): *Herausforderungen an die Informationswirtschaft : Informationsverdichtung, Informationsbewertung und Datenvisualisierung*. Proceedings des 5. Internationalen Symposiums für Informationswissenschaft (ISI '96), Humboldt-Universität zu Berlin, 17.–19. Oktober 1996 (pp. 25–31). Konstanz: UVK.
- DIN EN ISO 9241-210 (2010): Prozess zur Gestaltung gebrauchstauglicher interaktiver Systeme.
- Ellis, David, Cox, Deborah & Hall, Katherine (1993): A Comparison of the Information-seeking Patterns of Researchers in the Physical and Social Sciences. In: *Journal of Documentation*, 49 (4), 356–369. [doi:10.1108/eb026919](https://doi.org/10.1108/eb026919)
- Gantz, John F. (2008): The Diverse and Exploding Digital Universe. IDC white paper.
- Google Developers (2016): Google Maps APIs. <https://developers.google.com/maps/?hl=de> <08.12.2016>
- Kühnel, Andreas (2013): *Visual C# 2012 – das umfassende Handbuch* (6th ed.). Bonn: Galileo Press.

- Landwisch, P., Hemmje, M. & Fuhr, N. (2007): Ansatz zu einem konzeptionellen Modell für interaktive Information-Retrieval-Systeme mit Unterstützung von Informationsvisualisierung. In: Achim Oßwald, Maximilian Stempfhuber & Christian Wolff (Eds.): *Open Innovation : Neue Perspektiven im Kontext von Information und Wissen*. Beiträge des 10. Internationalen Symposiums für Informationswissenschaft und der 13. Jahrestagung der IuK-Initiative Wissenschaft, Köln, 30. Mai–1. Juni 2007 (pp. 327–332). Konstanz: UVK.
- Lange, Jürgen (2009): *Datenflut – Fluch oder Segen? Wie Sie mit Enterprise Search einfach und sicher Informationen finden. Ein strategisches Werkzeug für Unternehmen*. Frankfurt am Main: Frankfurter Allgemeine Buch.
- Longley, Paul (2011): *Geographic Information Systems & Science* (3rd ed.). Hoboken, NJ: Wiley.
- Loss, J. & Leitzmann, M. (2011): Ansätze zur verhältnisorientierten Adipositasprävention bei Kindern und Jugendlichen. In: *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz*, 54 (3), 281–289. [doi:10.1007/s00103-010-1232-7](https://doi.org/10.1007/s00103-010-1232-7)
- Macaulay, Catriona, Sloan, David, Jiang, Xinyi, Forbes, Paula, Loynton, Scott, Swedlow, Jason R. & Gregor, Peter (2009): Usability and User-Centered Design in Scientific Software Development. In: *IEEE Softw.*, 26 (1), 96–102. [doi:10.1109/ms.2009.27](https://doi.org/10.1109/ms.2009.27)
- Mazza, Riccardo (2009): *Introduction to Information Visualization*. Guildford, Surrey: Springer London.
- Minge, Michael & Thüring, Manfred (2009): Dynamics of User Experience. Judgments of Attractiveness, Usability, and Emotions Over Time. Technical Report 10-2009, Berlin: TU.
- Salton, Gerard & McGill, Michael J. (1983): *Introduction to modern information retrieval*. New York et al.: McGraw-Hill.
- Shneiderman, Ben (2003): The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In: B. B. Bederson & B. Shneiderman (Eds.): *The Craft of Information Visualization* (pp. 364–371). San Francisco: Morgan Kaufmann.
- Stephan, Veit (2016a): Mock-up Entwicklung von Food Environment Maps mit multiparametrischen, relationalen Daten (Dissertationsprojekt: Möglichkeiten der ortsbezogenen Visualisierung von komplexen, relationalen Daten am Beispiel von Food Environment Maps; Arbeitsbericht). Universität Regensburg.
- Stephan, Veit (2016b): Prototypengetriebene Evaluation von Visualisierungsmethoden bei kontextbezogenen LBAs (Dissertationsprojekt: Möglichkeiten der ortsbezogenen Visualisierung von komplexen, relationalen Daten am Beispiel von Food Environment Maps; Arbeitsbericht). Universität Regensburg.

Xu, Chen; Yang, Chaowei; Li, Jing; Xia, Jizhe; Qu, Xin; Sun, Min; Xu, Yan; Fay, Dan and Bambacus, Myra (2011): A Service Visualization Tool for Spatial Web Portal. In: *Proceedings of the 2nd International Conference on Computing for Geospatial Research & Applications, Washington, DC*. New York, NY: ACM, Article No. 31.

Yelp (2016): Yelp Dataset Challenge – Yelp. https://www.yelp.com/dataset_challenge <08.12.2016>

Doctoral Colloquium

**Session II – Information Retrieval,
Library Information Science, Social Media**

Relevance Clues

Developing an Experimental Design to Examine the Criteria Behind Relevance Judgments

Christiane Behnert

Hamburg University of Applied Sciences, Germany
christiane.behnert@haw-hamburg.de

Abstract

When users of information systems judge the relevance of search results, diverse criteria beyond topical relevance come into play. In this paper, we introduce the doctoral project *Relevance Clues*, which, through an experimental design, seeks to gain significant knowledge on the criteria by which users make relevance judgments.

Keywords: relevance; information behavior; interactive information retrieval; experiments; user criteria

1 Introduction

In modern library information systems (LIS), data as indicators for popularity (e.g., the number of clicks on a document) are available due to search engine technology. Web search engines consider such data as indicators for the relevance of a search result, i.e., they include these data in their relevance ranking algorithms. The project *LibRank – New Approaches to Relevance Ran-*

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 330–335.

ing in *Library Information Systems*¹ researched how such relevance factors can be adopted by libraries. Possible ranking factors suitable for LIS were identified (Behnert & Lewandowski, 2015) and systematically evaluated following standard procedures from Information Retrieval (IR) evaluation based on the methodological framework proposed by Tague-Sutcliffe (1992), which also, in the past, had been utilized for Web search evaluation using human relevance assessments (Lewandowski, 2012). Data as indicators for popularity were obtained from internal sources² (e.g., circulation data, the number of copies) and external sources (e.g., citation data), and were implemented in the test rankings in order to perform a total of three evaluation runs (Plassmeier et al., 2015). However, due to the research design and based on the assessment data, we do not know the criteria by which the jurors judged the documents.

At this point, the doctoral research project *Relevance Clues: Development and empirical examination of a model for relevance decisions on search results based on individual user criteria* (working title) sets in. The research goal is to gain significant knowledge on the criteria according to which users of academic search engines and LIS judge an information object (document surrogate) to be relevant to their individual information needs. In order to achieve this goal, an experimental research design will be applied.

2 Research questions

The project aims to answer the following research questions (RQs):

- (I) What clues within a surrogate do users use to judge on its relevance?
- (II) Which clues affect the relevance decision to what extent?
- (III) What influence does the use situation (e.g., the user's location, time pressure) have on the relevance judgment?
- (IV) What relevance criteria can be determined by the answers to RQ I–III, and how can they be weighed against each other?

1 www.librank.info

2 The data were obtained from *EconBiz*, an information portal for economics by the German National Library of Economics (www.econbiz.de).

The RQs are pursued in two steps: Based on an extensive literature review of current studies, a user model of relevance criteria will be developed and empirically examined through the conduction of a series of online experiments with human test persons. After reflecting the results, the user model will be altered accordingly.

3 Related research

3.1 Relevance criteria

In the 1990's, a shift from a system-oriented view towards a user-oriented view on relevance in Information Retrieval (IR) occurred, which is mirrored by studies on relevance criteria that were undertaken at this time (Mizzaro, 1997). In a recent publication, Saracevic (2016) synthesizes what we have learned from decades of relevance research. He provides an overview of 21 “observational, empirical, or experimental” studies on relevance clues, including the important works by Barry and Schamber (1998). They had analyzed criteria according to which users judge the relevance of a document. Among other studies, their results show that diverse criteria beyond topical relevance are involved in relevance judgments, for example, *validity*, *recency*, *availability*, and *credibility* of the information source.

With respect to the dynamic nature of the Web and its exponential growth, credibility and quality are both very important factors in order to filter and judge information retrieved by Web search engines (Rieh & Belkin, 1998). Credibility, in particular, can be considered in terms of *cognitive authority* (Wilson, 1983), which is highly subjective: A person is not only an expert but a cognitive authority as well, when his or her statements of knowledge are accepted by others as truth – the information is trustworthy – while he or she also influences other people's thoughts. Assessments of an author's cognitive authority are based on his or her present reputation and accomplishments (Rieh, 2009). Thus, in an academic context, information about an author's impact is helpful for users making relevance judgments. While general information about a document is presented as a surrogate, academic search systems integrate additional data into search results presenta-

tion. For example, Google Scholar's results include the number of citations or other versions of the particular work.

3.2 Document representations as objects of investigation

Documents have been represented by LIS using metadata ever since. Due to integrating electronic materials into modern systems and (1) full-text indexing as well as (2) enhancement with external data, for example, tables of contents and abstracts, users today are provided with a large amount of information about a document's content that they can utilize to judge its relevance. Further, Web 2.0 functionalities enable user comments and recommendations that can also be involved in the relevance decision process, as they indicate popularity. In academic information-seeking, popularity data would also include citation information, as mentioned above. So far, no studies on relevance assessments of surrogates including such popularity data and, at the same time, considering the user perspective have been published.

4 Methods

In order to examine which clues of a surrogate affect users' relevance judgments, a series of online experiments will be conducted. The essential characteristics of an experiment, as common in the field of psychology, are *manipulation* and *control*, which are the basic requirements for testing whether a causal relationship exists between a stimulus (independent variable) and an effect (dependent variable) (Sedlmeier & Renkewitz, 2007: 124–127). In relevance research so far, experimental research designs have not been applied very often, but this method follows the relatively recent trend in Interactive IR (IIR) that puts the user interacting with the system and the information objects in the center of retrieval evaluation (Kelly, 2009). Experiments are not only used for IIR evaluation but also to gain knowledge on user information behavior.

Within this project, the experiments will be conducted with students or academic staff using an online software tool. Subjects will be presented with a sequence of search results pages containing ten surrogates to a search query or information need in a randomized order. There will be two levels of poten-

tial relevance clues (of each independent variable) that will be manipulated (e.g., a low number of citations versus a large number of citations) to measure the effect on the relevance judgments (dependent variable). The clues to be tested will be selected based on the results of the literature review and a pre-test. Since it is assumed that there are diverse clues that affect relevance judgments, a multifactorial within-subjects design has to be developed. This requires a relatively large sample size. The goal is to recruit at least 400 subjects.

Acknowledgements

The research project presented in this paper is funded by a three-year Ph.D. scholarship from the Hamburg University of Applied Sciences, Germany. The doctoral candidate acknowledges her supervisors Prof. Dr. Ulrike Spree, Hamburg University of Applied Sciences, Germany, and Prof. Dr. Joachim Griesbaum, University of Hildesheim, Germany.

References

- Barry, C. L. & Schamber, L. (1998): Users' criteria for relevance evaluation: A cross-situational comparison. In: *Information Processing & Management*, 34 (2–3), 219–236. [https://doi.org/10.1016/S0306-4573\(97\)00078-2](https://doi.org/10.1016/S0306-4573(97)00078-2)
- Behnert, C. & Lewandowski, D. (2015): Ranking search results in library information systems – Considering ranking approaches adapted from web search engines. In: *The Journal of Academic Librarianship*, 41 (6), 725–735. <https://doi.org/10.1016/j.acalib.2015.07.010>
- Kelly, D. (2009): Methods for evaluating interactive information retrieval systems with users. In: *Foundations and Trends® in Information Retrieval*, 3 (1–2). <https://doi.org/10.1561/1500000012>
- Lewandowski, D. (2012): A framework for evaluating the retrieval effectiveness of search engines. In: C. Jouis, I. Biskri, J.-G. Ganascia, & M. Roux (Eds.): *Next Generation Search Engines* (pp. 456–479). Hershey, PA: IGI Global. <https://doi.org/10.4018/978-1-4666-0330-1.ch020>
- Mizzaro, S. (1997): Relevance: The whole history. In: *Journal of the American Society for Information Science*, 48 (9), 810–832. [https://doi.org/10.1002/\(SICI\)1097-4571\(199709\)48:9%3c810::AID-ASI6%3e3.0.CO;2-U](https://doi.org/10.1002/(SICI)1097-4571(199709)48:9%3c810::AID-ASI6%3e3.0.CO;2-U)

- Plassmeier, K., Borst, T., Behnert, C. & Lewandowski, D. (2015): Evaluating popularity data for relevance ranking in library information systems. In: *Proceedings of the 78th ASIS&T Annual Meeting*. <https://www.asist.org/files/meetings/am15/proceedings/submissions/posters/270poster.pdf>
- Rieh, S. Y. (2009): Credibility and cognitive authority of information. In: *Encyclopedia of Library and Information Sciences* (3rd ed.) (pp. 1337–1344). CRC Press. <https://doi.org/10.1081/E-ELIS3-120044103>
- Rieh, S. Y. & Belkin, N. J. (1998): Understanding judgment of information quality and cognitive authority in the WWW. In: *Proceedings of the 61st ASIS Annual Meeting* (pp. 279–289).
- Saracevic, T. (2016): *The notion of relevance in Information Science: Everybody knows what relevance is. But, what is it really?* Morgan & Claypool. <https://doi.org/10.2200/S00723ED1V01Y201607ICR050>
- Sedlmeier, P. & Renkewitz, F. (2007): *Forschungsmethoden und Statistik in der Psychologie*. München, Boston et al.: Pearson Studium.
- Tague-Sutcliffe, J. (1992): The pragmatics of information retrieval experimentation, revisited. In: *Information Processing & Management*, 28 (4), 467–490. [https://doi.org/10.1016/0306-4573\(92\)90005-K](https://doi.org/10.1016/0306-4573(92)90005-K)
- Wilson, P. (1983): *Second-hand knowledge: An inquiry into cognitive authority*. Westport, Conn.; London: Greenwood Press.

E-Publishing in Small Nations of the European Union

A Comparative Approach to Policy Frameworks and SME¹ Business Strategies

Anna Klamet

Edinburgh Napier University, UK
A.Klamet@napie.ac.uk

Abstract

The publishing industry is in a fundamental transition phase initiated by digitization. This transition involves production processes, communication options, and of course the ‘products’ themselves. While challenging the traditional assumptions, processes and structures of the book industry, e-books also offer a wealth of new opportunities to producers and readers: lower production and distribution costs, shorter time to market, earlier break-even points, easier vertical marketing strategies that target specific groups, new content features, and social reading as a product feature as well as a means for marketing and content discovery.

E-book distribution in most markets of the EU is dominated in particular by an oligopoly of international hardware and software companies as well as digital content distributors with their fenced-off e-book ecosystems. These ecosystems do not only hamper the differentiated bricks-and-mortar book-selling infrastructure and the cultural diversity in Europe, they also restrict the scope of action and the innovative potential especially of small-nation publishers – ‘small’ in this context relates among other things to the nation’s

¹ small and medium-sized enterprise

linguistic and economic status. Thus forcing publishers to think anew about innovative business strategies.

For this reason, this research project focuses on a theoretical investigation of potential business models for e-publishing, and their relationship to publishing operations, as well as on an empirical analysis of current policy frameworks and e-publishing practices within a number of small nations in the EU.

Keywords: publishing; publishing industry; digitization; small nations; SME; digital business strategies

1 Research plan

1.1 Research abstract

In very general terms, this research looks at the impact of digitization on small and medium-sized publishing houses focusing on trade publishing in what can be defined as small nations in Europe.² Due to their peripheral role in the European fiction market, these nations are often overlooked by researchers. That is why this study will focus on such small nations with the aim to better understand the struggles they face and will also address the question whether those small-nation publishers can benefit from e-publishing and if e-publishing could help to promote and protect Europe's cultural and linguistic diversity.

Technology has always been a major driver for change and innovation and especially digitization has affected almost every industry in the last decade. Hence, digitization and the spread of the internet have initiated a transition phase in publishing. This transition affects not only the production process but also the way of communication and the product itself as well as its distribution. The novel products offer new opportunities not only to readers but also to the industry itself: lower production and distribution costs, shorter time to market and therefore earlier break-even points. Online communities and social reading change the way books are consumed and social media

² The concept of nation within the scope of this research project can be defined as social groups that share a culture, history, ethnicity and language. Besides, small is a relative concept manifesting itself in categories such as geographical spread as well as in its linguistic and economic status (Boswell, 2014: 19).

allows for easier vertical marketing that targets specific groups. Furthermore, new features (enhancements) open new possibilities for authors and the way stories can be told. Therefore, digitization can be an advantage for small publishers: direct sales opportunities, new forms of marketing, cheaper production processes and so forth.

On the other hand, some aspects of digitization also pose a challenge for traditional publishers. For instance, self-publishing is becoming more and more popular which increases the number of e-books available online and making it difficult for text to stand out. This is especially true for English language titles. Moreover, e-book pricing is an issue: with lots of low-priced self-published books available, publishers are struggling to define a consistent pricing strategy. This is enforced through value added tax regulations in Europe which prohibit the application of a reduced rate on e-books. In addition, different VAT rates exist throughout Europe, creating an uneven field of play that privileges global media companies which can take advantage of these varying practices in order to reduce tax for their customers. Smaller actors are therefore more hesitant to invest in digital strategies.

Another important thing to keep in mind is that most markets of small nations in Europe today are largely influenced by an oligopoly of international hard- and software companies as well as digital content distributors which dominate the distribution of e-books with their mostly fenced-off e-book ecosystems through non-interoperable formats. Furthermore, they exercise censorship through prohibiting certain kinds of content to be made available on their platforms. This hampers the cultural diversity in Europe as well as restricts the scope of action and the innovative potential especially of publishers in small nations. So an important question is if there are ways to support the growth of a decentralised e-publishing sector in Europe.

To sum this up, the evolution of e-books in European book markets is highly diverse both in terms of market penetration, significant differences in policy debates, expectations by various stakeholders, and overall market context such as market size and cultural choices. Nevertheless, digitization is also influencing publishers in smaller nations in Europe as the linear supply chain has been transformed into a global network with powerful online retailers pushing publishers to establish a digital strategy. Furthermore, self-publishing, *Amazon*, *Apple* or *Google* all aim at replacing what used to be the core business of publishing. These developments, in an environment where technology and content merge together, are the reason why traditional publishing houses need to think about how they can benefit from the technologi-

cal change and if so, whether they need to innovate their publishing strategies and business models.

This notion is discussed in many recent newspaper and journal articles. But at the same time there is very little research that focuses on how the situation looks like in different markets at the moment and more specifically how small publishing houses are affected and how this is visible in new business models and business strategies. This is surprising as it can be assumed that small markets and hence small publishers are especially affected by the distorted competitive conditions. With this thesis this research gap shall be closed.

1.2 Research objectives

This research project aims to investigate how the business landscape of book publishing is changing with the increase in technological advancements and changing consumer behaviours. Of interest is, how the transformation in the publishing environment affects established publishing houses and what measures these publishing houses take in order to adapt and thus create sustainable businesses for the years to come in the digital age.

Most research on digitization in book publishing focuses on technology, organisational changes and product innovation but there is hardly any research that applies economic theories to the publishing industry. This is why this research uses the business model concept to analyse how and why established firms change their business practices and how this relates to their long-term strategy as well as how the context is influencing the options of publishing houses to innovate their business models. This will enable a new discussion about the change in publishing businesses.

By comparing and examining digital publishing practices of trade book publishers in three case studies, this thesis will answer questions about change and why some things stay the same. It is also a project about innovation in business strategies and change in business models as well as necessary adaptations in policy frameworks to be able to support publishing houses in their digital strategies. This leads to the following research questions, which will be answered using the methods described in the following chapter:

1. What are current and potential **innovative business models for e-publishing in small nations** in Europe and are they sustainable?
2. Which **internal and external factors** hinder or improve the **development** of innovative business models?

3. Does e-publishing help small-nation publishers to **reach a wider audience** and how do **oligopoly structures in the e-publishing distribution** hamper the development of the e-publishing sector?
4. How do current **policy frameworks for e-publishing** in those small nations and on a European level look like?

Through answering this research questions, this thesis will not only contribute to the theory of business models in book publishing but will also enhance the understanding of the future development of a decentralised e-publishing sector in Europe and lead to more informed policy-making. It also likes to contribute to managerial practice by means of supporting small-nation publishers with regard to their decision making. This involves organisational choices, program planning, technical specifications as well as revenue models – so that they can use the opportunities which digitization offers.

1.3 Overview of methodology

One part of the research will focus on the effects of digitization on publishing practices and especially on business models of small-nation publishers in Europe. A qualitative social research approach suits this topic best as there is little known about e-publishing practices in small nations and especially on the effects of digitization on business models. Qualitative research displays an explorative character in comparison to quantitative research and is therefore most appropriate in this context (Lamnek, 2010). The primary aim of qualitative research is to understand the social reality. In doing so, the real-life context is described initially followed by a deduction of hypotheses and theories from the collected data (Hermanns, 1992; Lamnek, 2010). Thus, the complexity of social problems and of the study object can be considered as well as the subjective meaning of the objective actuality for human beings. This subjective meaning is responsible for the actions of individuals and is therefore crucial for the understanding of the social reality (Flick, 2005; Lamnek, 2010).

In addition, I propose a case study design to gain a first overview of the current situation of e-publishing in small nations as this design is suited best when there is hardly any information about the object under examination (Eisenhardt, 2006). This research will mainly follow the stages of case study design as proposed by Eisenhardt so as to formulate hypotheses: definition of research questions, selecting subject matter, data collection, data procession, analysis, and documentation (ibid.: 219). Furthermore, I intend to

undertake a multiple-case design to make the predications more significant (Yin, 2003) and generalizable (Mayring, 2015). Through differences between the cases, correlations amid different characteristic values can be found (Gläser & Laudel, 2010). Moreover, comparisons are a vital part of the development of inductive generalisations as well as being central to procedures in grounded theory (Brewer, 2003). Empirical generalisation can be made by comparative research as the same process is studied in different fields and also leads to a better understanding of local specificities or generalities (ibid.). This is of relevance because of the differences in book culture and history as well as the different political and economic situations. Deductive comparative studies hereby tend to seek patterns of convergence between nations to support validity of general theory that is applied to understand and explain the social process under study (ibid.). Inductive comparative studies however tend to focus on patterns of variance to support the argument that social life is historically specific and culturally bound (ibid.).

For the purpose of data collection, expert interviews (semi-structured in-depth interviews) with relevant individuals in the publishing field will be held and then analysed with the help of a qualitative content analysis (Gläser & Laudel, 2010; Mayring, 2015). In order to understand why publishing houses engage in e-publishing or why not and how they intend to create value with new business models, interviews with managing directors of publishing houses as well as with employees responsible for e-publishing and innovation in these publishing houses seem most promising. The “advantage of this method [...] is that it enables you to get inside organizations and get a feel for how they work, allows you to explore issues in depth and helps you to see the world from the viewpoint of particular individuals located at particular points within the field” (Thompson, 2010: 406–407). Moreover, to understand the state of e-publishing in the respective nations, start-ups offering services which interfere with business models of publishers must be studied along self-publishing activities.

To analyse the impact of policy frameworks on the development of a successful e-publishing sector in these small nations, laws and discussions surrounding them need to be examined on an international, European and national level. Therefore, grounded theory seems to be a promising approach and shall therefore be analysed in detail on its applicability for the research project.

During the last months’ quantitative data on the publishing industry in the countries under examination, has been collected. This data will be used later

for triangulation during the analysis of the interview data. Furthermore, a questionnaire was designed on the basis of the research gap identified in the literature review and with the intention to answer the research questions.

References

- Boswell, Daniel (2014): *Publishing and the Industrial Dynamics of Biblio-cultural Identity in Catalan and Scottish Literary Fields*. PhD thesis, Edinburgh Napier University.
- Brewer, John (Ed.) (2003): *The A–Z of Social Research*. London: Sage
- Eisenhardt, Kathleen M. (2006): Building theories from case study research. In: P. Johnson & M. Clark (Eds.): *Business and Management Research Methodologies*, Vol. 4 (pp. 218–240). London: Sage.
- Flick, Uwe (2005): *Qualitative Sozialforschung: Eine Einführung* (3rd ed.). Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verlag.
- Gläser, Jochen, and Grit Laudel (2010): *Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen* (4th ed.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Hermanns, Harry (1992): Die Auswertung narrativer Interviews. Ein Beispiel für qualitative Verfahren. In: J. H. P. Hoffmeyer-Zlotnik (Ed.): *Analyse verbaler Daten: Über den Umgang mit qualitativen Daten* (pp. 110–141). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lamnek, Siegfried (2010): *Qualitative Sozialforschung: Lehrbuch* (5th ed.). Weinheim: Beltz.
- Mayring, Philipp (2015): *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12th ed.). Weinheim: Beltz.
- Thompson, John B. (2010): *Merchants of Culture: The Publishing Business in the Twenty-First Century*. Cambridge: Polity Press.
- Yin, Robert K. (2003): *Applications of case study research* (2nd ed.). London: Sage.

How Can Social Software Enhance Training Measures within Software Implementations in the Financial Sector?

Anthea Moravánszky

Institute of Information Science and Language Technology
University of Hildesheim, Germany
anthea.moravanszky@uni-hildesheim.de

Abstract

This paper summarizes a doctoral thesis project which argues to adopt End-User Training (EUT) as an acceptance enhancing factor of a software deployment. The main research object of this project is to explore how social software can be used in combination with, or as a substitute to traditional and well-established training methods for EUT within a software implementation. The target group will be within the financial sector. The aim is thus to build a bridge between research on the acceptance of (organizational) software implementation, End-User Training and learning research from the perspective of information science. Even if the target group of this project is limited to the financial sector it should be possible to apply the presented design dimensions to other sectors.

Keywords: social software; End-User Training; EUT; software training; software acceptance; software deployment

1 Introduction

A software implementation has several factors which could influence its success. Although the technical deployment plays an important role, ultimate

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 343–347.

success is determined by non-technical aspects (Kohnke & Bungard, 2005: 41). One of these aspects is user acceptance which can be significantly enhanced through suitable trainings (Kohnke, 2015: 420–445). In addition to many other fields of application, also the training sector has been promised major changes with web 2.0 and social software. Learning and working should be even more online and mobile.

The term “learning 2.0” was created in accordance with the concept scheme “web 2.0” (Stoller-Schai, 2011). Both terms strive to highlight the possibilities for participation and or the interactivity of the users or consumers in certain areas. However, learning 2.0 is “neither an already established canon of topics and methods, nor does it mean that established and proven topics and methods are to be completely replaced. Instead of a revolutionary change, it is rather an evolutionary development, an enhancement of the tried and tested with new aspects” (ibid.: 10). And these enhancements are “web-based applications and services that are used for social interactions within groups that support human communication and cooperation, thus sharing the structure and maintenance of social networks, as well as publication and distribution of information within social networks” (Manouchehri, 2010: 1) – in short, “social software”.

Did learning-revolutionary development happen already in companies? To what extent do companies already use social software in the training context and more specifically, in the End-User Training of software? And what kind of social software is actually suitable for enhancing or even replacing established End-User Training methods? Manouchehri (2010), Back (2012) and others have already shown some potential benefits and applications of social software in companies. Social software offers

“new forms of direct, personal communication, new possibilities in the creation and use of collective intelligence and community performance, an effective way of dedicated provision, preparation and exploitation of already existing knowledge, visualization of communication, interaction and knowledge as well as new ways in the search for information, common resources or even knowledge sources” (Manouchehri, 2010: 68).

How can these benefit potentials be used specifically to enhance End-User Training? The answer to this question is the object of investigation of this interdisciplinary doctoral thesis project.

2 Research questions

The overall aim of the doctoral thesis project is to highlight the benefits that using social software within End-User Trainings of software can grant for (financial) companies. The research questions which frame this project are as follows:

RQ1: What kind of social software is suitable for EUT?

- (a) Which social software applications are applicable to the leaning/training context?
- (b) How can they be combined with traditional EUT methods?
- (c) How can they replace traditional EUT methods?

RQ2: How is social software already used in the training context in (financial) companies?

RQ3: What are the benefits for companies to use social software within End-User Trainings?

3 Approach and methodology

An iterative approach is used in the empirical study. In a first step, based on the results of the literature research and its derived desiderata an inventory of the as-is condition is made. For this purpose, employees from the financial sector and educational experts (those working in financial institutions as well as external experts) are interviewed. In order to obtain a comprehensive picture, a combination of qualitative and quantitative examination instruments should be used.

To collect as much data as possible the employee's information should be gathered with an online questionnaire. At the same time, qualitative, guided expert interviews should be performed with two central target groups: company-based trainers, who are active in a financial enterprise and external experts. In order to benefit from the findings of the previous interviews and to illuminate certain new aspects from several perspectives, the interview guide will be consistently checked for possible extensions after each interview. The findings of this first iteration will be matched and combined with the desiderata of the literature research. The outcome of this will be a training matrix.

The levels of the matrix contain the social software applications which have been identified as appropriate for the training context, but also traditional EUT methods, and recommendations about suitable combinations and influence categories (for example organizational framework, training goal and/or content). This is followed by a feedback session with the training experts working in the financial sector. The content of this iteration step is a result presentation of the employee survey and its determined discrepancies with the results of the expert interviews. A concept test of the training matrix should be performed within the feedback sessions as well which might result in further improvements of the training matrix.

In at least two financial companies an actual implementation of the training matrix is to be carried out. Appropriate elements suitable for the selected software introduction are selected from the training matrix and implemented and trained. Afterwards there should be an employee survey done to measure the success. If possible, usage figures and hotline reports should also be included in the success measurement. The last iteration step, which completes the validity test of the training matrix, is a further feedback session with the training experts. This could either be a group discussion or even a symposium.

4 Status quo and next steps

This doctoral thesis project has just started in autumn 2016. While literature research and deriving of desiderata is done now, the next planned step is the start of the empirical study in spring 2017.

References

- Back, A., Gronau, N. & Tochtermann, K. (Eds.) (2012): *Web 2.0 und Social Media in der Unternehmenspraxis. Grundlagen, Anwendungen und Methoden mit zahlreichen Fallstudien*. München: Oldenbourg
- Kohnke, O. & Bungard, W. (2005): *SAP-Einführung mit Change Management : Konzepte, Erfahrungen und Gestaltungsempfehlungen*. Wiesbaden: Gabler

- Kohnke, O. (2015): *Anwenderakzeptanz unternehmensweiter Standardsoftware: Theorie, Einflussfaktoren und Handlungsempfehlungen*. Wiesbaden: Springer.
- Manouchehri Far, S. (2010): *Social Software in Unternehmen: Nutzenpotentiale und Adoption in der innerbetrieblichen Zusammenarbeit*. Köln: Eul.
- Stoller-Schai, D. (2011): Lernen 2.0 in der Finanzbranche. In: Armin Trost and Thomas Jenewein (Eds.): *Personalentwicklung 2.0: Lernen, Wissensaustausch und Talentförderung der nächsten Generation* (pp. 109–129). Köln: Luchterhand.

Enriching a Library Portal (*EconBiz*) with Altmetrics

Kaltrina Nuredini

ZBW – Leibniz Information Centre for Economics, Kiel, Germany
k.nuredini@zbw.eu

Abstract

In this thesis, the opportunities for introducing altmetrics in library information systems are explored. We discuss studies looking at altmetrics data for Business and Economics literature. The focus of the thesis is to highlight the relevant altmetric indicators for journals in a specified discipline, their aggregation level and visualization modes. The most important findings are discussed and we present some preliminary suggestions for future integration of altmetrics data in library information systems.

Keywords: social media; altmetrics; library information systems

1 Motivation

Social media-based bibliometric indicators, so called *altmetrics*, can add an alternative filtering layer to library collections especially to libraries with scholarly focus. Altmetrics can help economic researcher evaluate the impact of the articles they want to read or journals where they want to publish (Nuredini & Peters, 2015, 2016) for free and in a very short time period. However, there is still room for investigating the list of relevant indicators by deeming to understand how they should be presented in library collections. Moreover, questions like what type of aggregation of altmetrics data is ade-

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 348–354.

quate should be answered (aggregation on provider level or indicator level). *Altmetric.com* for example aggregates altmetric indicators from various sources in an altmetric score.¹ This methodology avoids users' information overflow when listing many indicators scores separately.

Specifically, the aim of this proposal is to answer the following research question: *RQ. To what extent can altmetric information enrich libraries with a specified focus?* Additionally, the proper modes of visualizations for presenting altmetric indicators for journals in *EconBiz*² will be examined. *EconBiz* is an online portal with a focus on Economic and Business Studies literature. Visualization modes or filters will act as a proxy to help users filter publications based on the accurate altmetric scores on journal level. Adding answers to the research question mentioned above, two studies in this field have been made so far. However, more future studies should be done to complete the investigation.

2 Introduction

Technology is present everywhere and its vitality is indisputably affecting the nature of libraries. By embracing new technology, libraries are no more places with card files and books, they are fueled with digital revolution, research and especially with a new influence from social media services. With respect to the introduction of social media in libraries, they are increasingly interested in new measures which can help in research assessment, showcase the performance of institution's scholarly output, and increase authors' awareness of their research impact (NISO, 2016) that might complement the ability of traditional measures. New measures known as altmetrics use sources from the social web such as *Twitter*, *Facebook*, news or reference management tools to quantify the impact of scholarly publications on social media users (Priem et al., 2010). Concerning the benefits of altmetric indicators, information infrastructure providers and libraries have increasingly become interested in using altmetrics data to facilitate filtering of publications, pro-

1 <https://help.altmetric.com/support/solutions/articles/6000060969-how-is-the-altmetric-attention-score-calculated->

2 www.econbiz.de

viding context information to publications, and help patrons - as well as library staff – in assessing the relevance of publications. Moreover, publishing houses and aggregators of altmetric data popularize social media indicators by attaching them to their products and along with it promoting those articles or other research outputs. By now, social media indicators are close to being ubiquitous in scholarly communication environments. This perceived prevalence of altmetrics and its usage as a scientometric tool, oftentimes sold as easy to understand and easy to implement (e.g., by bookmarklets³), can blur the pitfalls of such approach. Hence, especially independent and non-profit institutions like libraries need to know, for example, which aspects can be implemented in a reasonable way, where sufficient data is available for valid analyses, what altmetrics window (analogous to the citation window) should be used, and which altmetrics aggregator is the best choice for the goals set. It is an ongoing debate what altmetrics show, exactly. Nevertheless, different studies found that altmetrics are the complements of the traditional indicators for research evaluation (Bar-Ilan et al., 2012); and Loach and Evans (2015) suggest that altmetrics reflect a different type of impact. Haustein (2016) highlights that any metric whether it is citation or social media based, has to be wisely chosen depending on the assessment aim. Bornmann (2014) argues that altmetrics offer four benefits in impact measurement. First, altmetrics offer broader access to the opinions of a wider audience for research articles besides citation metrics used only by scientific authors. Similarly, they do not only allow evaluation of scholarly publications but can also be applied to a diversity of products such as presentation slides, algorithms, software applications etc. Additionally, altmetrics can speed up impact evaluations of publications by showing online attraction just a few days or weeks after their publication date. And lastly, by not relying on a single provider of citation counts but the web APIs of widely used social media platforms free access to the altmetrics data is possible which facilitates its analysis and interpretation. All these benefits are ignored by most of the traditional indicators.

A lot of research has been done that studies altmetrics from a multidisciplinary perspective such as in Costas et al. (2014), Alhoori et al. (2014) but also a very small number of studies about altmetrics and its implications in Economic and Business Studies journals. In Nuredini and Peters (2015) it has been mentioned that economic researchers deal with many alternative publications formats and they find it difficult to decide what is important for them

3 <https://www.altmetric.com/products/free-tools/bookmarklet>

to read or where research findings should be present. Thus, altmetric data in the field of Economics and Business Studies have been collected to exploit the importance of altmetric indicators and add a further layer to traditional research evaluation in this field.

3 Results

With two studies done so far (Nuredini and Peters, 2015, 2016) we found out the relevant altmetric indicators for Economics and Business Studies Journals based on two the altmetric data providers *Mendeley*⁴ and *Altmetric.com*⁵.

From our data analysis, we showed that *Mendeley* can be suggested as a relevant source for journals in Economic and Business Studies because of the high coverage of journal articles found. According to Wouters et al. (2015) *Mendeley* readership counts appear to be the most promising altmetric indicator because of two reasons: 1) easy use of automatic data collection and 2) huge coverage of articles of different fields. Fairclough and Thelwall (2015) have similar findings and claim that *Mendeley* reader counts are free and may be useful for those who can't access citation databases. *Mendeley* user readership information can act as a support to choose the right articles for reading (Nuredini & Peters, 2015). Based on several studies it can be concluded that *Mendeley* covers readers (users) which fall in the category of younger researchers i.e., Bachelor, Master, PhD students (Fairclough & Thelwall, 2015; Nuredini & Peters, 2015).

Additionally, our observation with *Altmetric.com* shows that altmetric data scores make more sense for recently published articles because they appear to be mentioned more often online than earlier articles (Nuredini & Peters, 2016). Alhoori and Furuta (2014) acknowledge that altmetrics have the potential to predict delayed citation-based metrics. Likewise, we have learned that altmetric data scores are not fully related with citation counts. Moreover, altmetrics and citation counts indicate a positive but low correlation on article level which is confirmed by Alhoori and Furuta as well. On the other hand, journal level altmetrics have a moderate correlation with citation

4 <https://www.mendeley.com/>

5 <https://www.altmetric.com/>

counts and H-index values but weak correlations with other citation based metrics (ibid.). The correlation between impact factor (IF) and Altmetric Score on journal level is low but positive (Spearman $r = 0.314$ and Pearson $p = 0.169$) – hence we can conclude that articles from highly cited journals are not receiving substantial attention online. Out of 8 altmetric sources provided by *Altmetric.com* our study results confirm that 4 of them (*Twitter*, *Blogs*, *News* and *Facebook*) are the places where paper of economists are mostly found there making them good candidates as relevant altmetric indicators. Additional findings from our study show that articles got the highest coverage in *Twitter* with 88% whereas Hamerfelt (2014) confirms comparable results by stating *Twitter* as a source with highest coverage of 20% for his data set in the humanities. He also suggests that *Twitter* might be an alternative indicator for measuring the impact of books because of the wider audience in this platform whereas *Mendeley* covers more scholarly publication formats.

From our analysis we can see that for articles in Economic and Business Studies altmetrics data is still rather sparse, although availability has increased for more recent articles. Therefore, higher aggregation levels such as journal level, may overcome the sparsity of altmetrics data. By doing so, it will be ensured that for every record altmetric information could be displayed which lowers, or even avoids, user frustration. This statement might hold for *Altmetric.com* sources. Altmetric indicators would be an alternative to traditional bibliometric methods which will help analyzing and measuring the impact of research also for Economic and Business Studies journals (Nuredini & Peters, 2016; Hamerfelt, 2014). Altmetric indicators so far have been explored for 30 journal articles in these fields only. In future studies we would like to explore all journals listed in the *Handelsblatt* ranking⁶ in both fields of Economics and Business Studies, with more than 1000 journals.

These journals will be queried in *Mendeley* and *Altmetric.com* for altmetric information and we want to determine the best altmetric aggregation for these sources. After finding the relevant altmetric information for the different aggregation levels, the next step will be to find out what data can be visualized and attach this visualization model to the *EconBiz* data collections.

6 <http://tool.handelsblatt.com/tabelle/?id=33&so=1a&pc=25&po=1225>

References

- Altmetrics for Librarians (2016): 100 + tips , tricks , and examples Brought to you by Altmetric. <http://doi.org/10.6084/m9.figshare.374983>
- Alhoori, H. & Furuta, R. (2014): Do altmetrics follow the crowd or does the crowd follow altmetrics? In: *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, pp. 375–378. <http://doi.org/10.1109/JCDL.2014.6970193>
- Bar-Ilan, J., Haustein, S., Peters, I., Priem, & J., Shema, H., et al. (2012): Beyond citations: Scholars' visibility on the social Web. In: *Proceedings of the 17th International Conference on Science and Technology Indicators*. Montréal, Canada, pp. 98–109.
- Bornmann, L. (2014): Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. <http://arxiv.org/abs/1406.7091>
- Costas, R., Zahedi, Z., Wouters, P. F. (2014): Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multi-disciplinary perspective. <https://openaccess.leidenuniv.nl/handle/1887/23041>
- Fairclough, R. & Thelwall, M. (2015): National research impact indicators from Mendeley readers. In: *Journal of Informetrics*, 9 (4), 845–859. <http://doi.org/10.1016/j.joi.2015.08.003>
- Hammarfelt, B. (2014): Using altmetrics for assessing research impact in the humanities. In: *Scientometrics*, 1419–1430. <http://doi.org/10.1007/s11192-014-1261-3>
- Haustein, S. (2016): Grand challenges in altmetrics: heterogeneity, data quality and dependencies. In: *Scientometrics*. [doi:10.1007/s11192-016-1910-9](https://doi.org/10.1007/s11192-016-1910-9)
- Library, B. H. (2009): Case Study. In: *American Libraries*, 40 (6/7), 97. <http://doi.org/Case Study>
- Loach, T. V. & Evans, T. S. (2015): Ranking Journals Using Altmetrics. arXiv: 1507.00451
- National Information Standards Organization (2016): Outputs of the NISO Alternative Assessment Metrics Project. http://www.niso.org/apps/group_public/download.php/17091/NISO_RP-25-2016_Outputs_of_the_NISO_Alternative_Assessment_Project.pdf
- Nuredini, K. & Peters, I. (2015): Economic and Business Studies Journals and Readership Information from Mendeley. In: F. Pehar, C. Schlögl & C. Wolff (Eds.): *Re:inventing Information Science in the Networked Society*. Proceedings of the 14th International Symposium on Information Science (ISI 2015), Zadar, Croatia (pp. 380–392). Glückstadt: Hülsbusch.

- Nuredini, K. & Peters, I. (2016): Enriching the knowledge of altmetrics studies by exploring social media metrics for Economic and Business Studies journals. In: *Proceedings of the 21st International Conference on Science and Technology Indicators (STI Conference 2016), València (Spain), September 14–16.*
- Peters, I., Jobmann, A., Eppelin, A., Hoffmann, C. P., Künne, S. & Wollnik-Korn, G. (2014): Altmetrics for large, multidisciplinary research groups: A case study of the Leibniz Association. In: *Proceedings of the Libraries in the Digital Age Conference, Zadar, Croatia.* <http://ozk.unizd.hr/proceedings/index.php/lida/article/view/162/138>
- Priem, J., Taraborelli, D., Groth, P. & Neylon, C. (2010): Altmetrics: A manifesto, 26 October 2010. <http://altmetrics.org/manifesto>
- Wouters, P., Thelwall, M., Kousha, K., Waltman, L., De Rijcke, S., Rushforth, A. & Franssen, T. (2015): *The Metric Tide: Literature Review* (Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management). <http://doi.org/10.13140/RG.2.1.5066.3520>

Poster Presentations

An Information System for the Analysis of Color Distributions in *MovieBarcodes*

*Manuel Burghardt*¹, *Katharina Hafner*², *Laura Edel*²,
*Sabrin-Leila Kenaan*², *Christian Wolff*¹

Media Informatics Group, University of Regensburg, Germany

¹{manuel.burghardt, christian.wolff}@ur.de

²{katharina.hafner, laura.edel, sabrin-leila.kenaan}@studi.uni-regensburg.de

Keywords: information systems; film studies; color analysis

We present an ongoing project from the field of quantitative film studies, sometimes also referred to as *Cinematics* (Tsivian, 2009). Most of the related work in this area is focused on quantitative analyses of shot lengths and distributions.¹ In this paper, we suggest *color* as an additional quantitative parameter for movie analysis and describe an information system that allows scholars to search for movies via their specific color distribution. As a source of condensed movie color information, we make use of the *MovieBarcode*² database. A *MovieBarcode* is created by skewing each frame of a movie to be only 1 pixel wide. Lining up all these frames in a row creates a barcode-like visualization of the most dominant colors in a movie (fig. 1).³ Our information system makes use of the *color diff*⁴ library to map more than 1,500

1 Cf. the extensive cinematics bibliography at <http://cinematics.lv/articles.php> (all URLs in this paper were last accessed on 15 November 2016).

2 available via <http://moviebarcode.tumblr.com>

3 For a similar visualization approach cf. Barbieri et al. (2001).

4 *Color diff* is an implementation of the CIEDE2000 color difference algorithm. Available via <https://github.com/markusn/color-diff>.

MovieBarcodes to a palette of 11 standard colors (cf. Welsch & Liebmann, 2006). In addition to these individual color profiles, we also collect metadata⁵ (genre, year, director, country, etc.) and keywords from the movies' subtitles.⁶

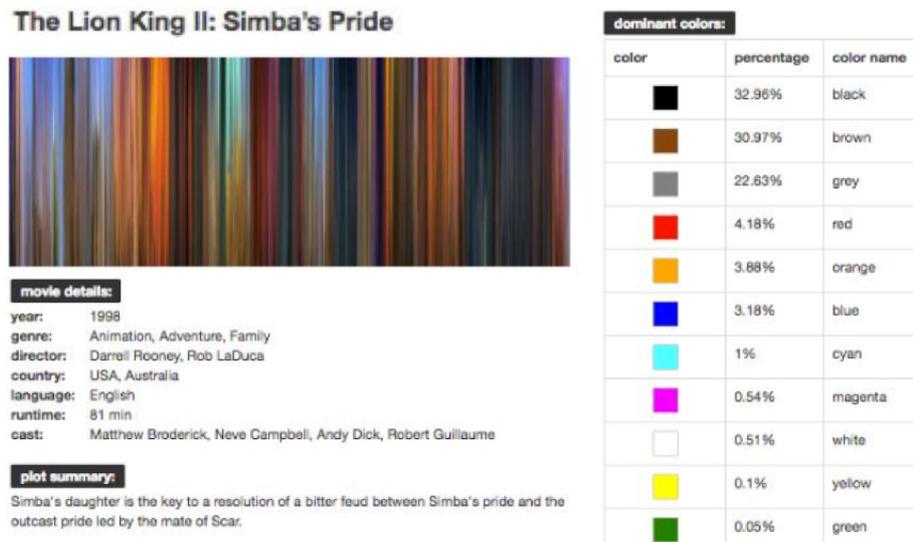


Fig. 1 Example of a *MovieBarcode* and its specific color distribution

The tool can be used to search for movies based on their color distributions, or to identify general trends in the use of color in specific genres or periods of time, or in combination with certain keywords. Example questions that can be answered with our information system are:

1. What is the most frequent color in horror movies as compared to comedies?
2. How did the use of color in movies develop from the 1940s to the 1980s?
3. What are the most frequent words in movies that contain a lot of blue?

We are currently testing the system with scholars from the film studies area. In its current implementation, our system can be used as a “distant watching” tool (cf. Howanitz, 2015), i.e. it is used for the generation of new research questions or to test early hypotheses by investigating a large collection of movies from a quantitative perspective. As a next step, we want to

⁵ The meta information is available via <http://www.imdb.com>.

⁶ Subtitles are available via <http://www.opensubtitles.org>.

extend the system to become a rich-prospect browser (Ruecker et al., 2011), i.e. the tool will allow scholars to zoom into specific movies and to investigate them on more detailed levels of analysis, ranging from single frames to shots and scenes.

References

- Barbieri, M., G. Mekenkamp, M. Ceccarelli and J. Nesvadba (2001): The Color Browser: A Content Driven Linear Video Browsing Tool. In: *Proceedings of the IEEE International Conference on Multimedia and Expo* (pp. 808–811).
- Howanitz, G. (2015): Distant Watching: Ein quantitativer Zugang zu YouTube-Videos. In: *Book of Abstracts, DHd 2015*.
- Ruecker, S., M. Radzikowska, and S. Sinclair (2011): *Visual Interface Design for Digital Cultural Heritage*. Farnham et al.: Ashgate Publishing.
- Tsivian, Y. (2009): Cinematics, Part of the Humanities' Cyberinfrastructure. In: M. Ross et al. (Eds.): *Digital Tools in Media Studies* (pp. 93–100). Bielefeld: Transcript.
- Welsch, N. and C. C. Liebmann (2006): *Farben. Natur, Technik, Kunst*. München: Elsevier.

Data Streams in *linked.swissbib.ch*

The Swiss Metacatalog in the Linked Open Data Cloud

Nicolas Prongué, René Schneider

University of Applied Sciences HEG Genève, Switzerland
{nicolas.prongue, rene.schneider}@hesge.ch

Keywords: bibliographic data; daily updates; data processing; data stream; interlinking; library metacatalog; linked data; swissbib

The project *linked.swissbib.ch* aims to integrate the metacatalog *swissbib* into Linked (Open) Data, by transforming, interlinking and enriching the data. This implies the setting up of an infrastructure providing on the one hand a data service for other applications via a specific interface and on the other hand an improved interface for the end user (e.g. a searcher). Expected benefits of that approach include better data interoperability, an easier data reuse and a more enriching user experience. The project is jointly carried out by the University of Applied Sciences HEG Genève, the University of Applied Sciences HTW Chur and the Basel University Library.

The aim of this poster is to describe the whole system infrastructure of *linked.swissbib.ch*, and particularly to reveal the challenges related to metadata operations at the level of a metacatalog, namely at a level where data is not produced but only harvested at a daily frequency. The originality of this poster lies in its representation of the data environment, using the metaphor of the data lake and data stream, introduced by Redman (2008).

Some fifteen Swiss library networks are providing bibliographic metadata to *swissbib*, which processes, deduplicates, and then transforms it. For these operations, various data formats are managed, among others the traditional library format MARC/XML, as well as more recent RDF based formats like

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 359–361.

RDF/XML, JSON-LD or NT. For the transformation process, the software Metafacture¹ is mainly used. As the data store of linked.swissbib.ch is actualised daily, it consists of a continually changing mass causing further challenges for the data management. On the one hand, the attribution of unique and permanent identifiers in the form of URIs – being the first of the four Linked Data grounding principles of Tim Berners-Lee (2010) – is problematic, because the records vary every day in function of the deduplication operation of swissbib. On the other hand, the interlinking with external datasets like VIAF or DBpedia – being the fourth of these principles – becomes tricky as it relies on the existing URIs. These data processing operations are made even more complex due to the large amount of data (about 21 million records, the equivalent of ca. 39 GB) composing the metacatalog. Once the data deduplicated, transformed and interlinked, it is made available for computer clients and human users. Firstly, massive data reuse is possible for machines through a RESTful API. Since the library networks providing swissbib have various terms of use, this could lead to difficulty of attribution for the re-user. To address this issue, a mechanism filters the data in such a way that only CC0-compatible records are made accessible via the API. Secondly, an experimental interface is being developed for the end user, whose goal is to offer an improved search and exploration experience based upon the new interconnected data.

Linked Open Data is often said to be a very promising technology. Nevertheless, its implementation into a concrete and sustainable application reveals extremely complex data processing operations. This is notably the case for a library metacatalog treating exclusively secondary data. This poster highlights the key challenges of such an approach and illustrates the solution found for the specific case of linked.swissbib.ch. The result will be of interest for all researchers who face similar problems in other metacatalogs.

Project website:

http://www.swissbib.org/wiki/index.php?title=Linked_swissbib

1 Software website: <https://github.com/culturegraph/metafacture-core> <28.12.2016>

References

Berners-Lee, T. (2010): Linked Data. <http://www.w3.org/DesignIssues/LinkedData.html>

Redman, T. C. (2008): *Data driven: profiting from your most important business asset*. Boston, Mass.: Harvard Business Press.

For You, Me, Them?

Choosing Designated Communities for Personal Digital Archiving

Achim Oßwald

Institute of Information Science, TH Köln, Germany
achim.osswald@th-koeln.de

Abstract

The paper addresses motivations for choosing designated communities when applying measures for archiving personal digital data and digital objects. Results of two case studies focusing on designated communities and the types of archived digital objects are reported and discussed.

Keywords: personal digital archiving; designated communities; motivation

Motivations and designated communities for personal digital archiving

In professional settings concepts for digital storage and preservation of data are pushed by funding organizations and other stakeholders. Management and preservation e.g. of research data have become a standard expectation not only in research environments but in the business area, too. More and more tools and concepts for dealing with data and files are available but there is still reluctance to apply them.

Digital objects and digital communication have gained relevance in private contexts, too. The number and volume of digital files created and stored by individuals has grown from year to year. Pictures, small movies and

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 362–364.

sound files are the fastest growing types of digital data. They are stored e.g. on smartphones, notebooks and by using cloud-based services. The availability of large and cheap storage media accelerates this trend. Most files are saved and stored automatically by software based on the names assigned to the files. Personal Digital Archiving (PDA) is a concept of digital curation and preservation helping individuals and groups to come to grips with their digital collections and to preserve content that is important to their lives in trustworthy, long-term ways.

In these private settings (adapted) professional concepts for selection as well as grouping or naming files or adding metadata are rarely used. Most of the tools and concepts developed for professional environments are not suitable because resources and knowledge are missing in private settings. As a consequence, special recommendations e.g. provided by Library of Congress (Library of Congress, 2016) and nestor Personal Digital Archiving working group (nestor AG PDA, 2016; Oßwald et al., 2016) are more applicable in this context. Nevertheless, their application is laborious. Special motivation is needed, which is partially drawn from the expected benefit, acknowledgement and perhaps gratitude of the people who will reuse a structured and well-named collection of text files, photos or other data. Among these designated communities there could be the owner of the files him- or herself. Motivation for curation activities might be very personal (Kim, 2013) or family-related (in a great number of cases; see e.g. recent presentations at PDA, 2016), in other cases based on other private links to different designated communities at present or in the future. Nevertheless, there is little knowledge about age-related differences in motivation of individuals putting effort in the organization, curation and preservation of personal digital data.

The poster presents results of two case studies with Library Science students (of TH Köln) and Cologne citizens who have taken part in a public workshop at Cologne Public Library (Digitale Werkstatt, 2016). Motivations of both groups are explored when applying basic activities of personal digital archiving in relation to the designated communities they have in mind.

Two hypotheses are explored: (1) The younger people are, the more are they themselves the designated community of their personal digital archiving activities. (2) Digital objects are curated and archived differently depending on age and family ties.

Acknowledgements

The author would like to thank those library science students of TH Köln who have contributed to his research by providing personal perspectives and preparing the public workshop on PDA at Cologne Public Library.

References

- Digitale Werkstatt (2016): Digitale Werkstatt – Cloud, Festplatte oder CD-ROM im Schuhkarton? Digital workshop – cloud, hard disk or CD-ROM in a shoe box? In: *Workshop at Cologne Public Library, Cologne, Germany, November 17th, 2016*.
- Kim, Sarah (2013): *Personal Digital Archives: Personals Documents, Preservation of Documents, Preservation of Self*. PhD thesis, Graduate School of the University of Texas, Austin, USA. <https://repositories.lib.utexas.edu/bitstream/handle/2152/21721/KIM-DISSERTATION-2013.pdf>
- Library of Congress (2016): Personal Archiving. <http://www.digitalpreservation.gov/personalarchiving/><23.10.2016>
- nestor AG PDA (2016): nestor AG Personal Digital Archiving. <https://wiki.dnb.de/display/NESTOR/AG+Personal+Digital+Archiving> <23.10.2016>
- Oßwald, Achim, Martin Iordanidis and Natascha Schumann (2016): Simplify your Digital Life: nestor-AG Personal Digital Archiving gegründet. In: *B.I.T.online* 19 (4), 439–441. <http://www.b-i-t-online.de/heft/2016-05-nachrichtenbeitrag-iordanidis.pdf>
- PDA 2016 (2016): Personal Digital Archiving 2016. University of Michigan Library, Michigan, USA. <http://www.lib.umich.edu/pda2016> <23.10.2016>

Grasping the Materializations of Practices in Digital Humanities

A Semantic Research Environment for Analyzing Exam Grading Practices in German High Schools

*Christoph Schindler¹, Julian Hocker¹, Lars Müller¹,
Maria Maleshkova², Tobias Weller²*

¹ DIPF, Germany

{schindler, julian.hocker, l.mueller}@dipf.de

² KIT, Germany

{maria.maleshkova, tobias.weller}@kit.edu

Keywords: digital humanities; research environment; practice theory; qualitative research; distant reading

1 Motivational background

Recently, practice theory has been established alongside humanities and social sciences in historical research and has aligned analysis according to materializations of loosely interlinked arrangements of practices, materialities, and actors (cf. Reh, 2014). While digital humanities have proven its great capacity to deal with massive data (distant reading), the digital enhancement of qualitative approaches is lacking (cf. Drucker, 2012). This desideratum is addressed in this project which uses a semantic graph to create a linked web and offering tools for combining qualitative and quantitative analysis.

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 365–367.

2 Research design and digital humanities

Empirically, the case study addresses a corpus of 2.000 digitized and around 800 transcribed German essays from high-school graduates (1882–1972). In the handwritten essays of the high school graduates, the teachers' reviews are materialized which need to be profoundly and expressively connected to sources of the educational and governmental apparatus (e.g. school protocols, pedagogical publications, regulations, legal provisions). The corpus encompasses essays from Berlin/Brandenburg, Baden-Wuerttemberg, and Bavaria allowing a comparison of arrangements of the different practices of grading. The selected research environment Semantic CorA¹ offers to semantically collect and query sources as well as researchers' comments within one single system. Thus it supports integration of qualitative and quantitative research approaches adjust the research project to its specifics and needs. Therefore, a participatory design approach with an agile development and rapid prototyping ensures involvement of researchers as well as archivists, a data librarian, computer scientist, and information scientists into system development.

3 First outcomes and future plans

A first main outcome is the establishment of a workflow for the research data pipeline, which needed to grasp the archive research, essay collection and curation (e.g. schools, archives), digitization, enrichment, and transcriptions and the detailed semantic annotations in the research environment as well as a sustainable solution for re-using the data. Besides the data/metadata translations (METS/MODS, TEI, Wiki-Semantics), we realized a collaborative ontology development for specifying the research objectives (actors, materialities, practices) balancing complexities (e.g. fluidity of geographic boundaries, various educational apparatuses) to create a semantic graph, whereby an openness of formalizing entities and properties needed to be considered. Software development will focus on a tool for annotation and as well as on

1 http://semantic-cora.org/index.php/Main_Page

a tool for analyzing text within Semantic CorA (cf. Rockwell & Sinclair, 2016).

Acknowledgements

The project is funded by the Leibniz Competition 2016 entitled “Abiturprüfungspraxis und Abituraufsatz 1882 bis 1972”. It is realized in cooperation of the DIPF (BBF, IZB), the AIFB at the Karlsruhe Institute of Technology, and the department of German Literature at the HU Berlin.

References

- Drucker, J. (2012): Humanistic Theory and Digital Scholarship. In: Matthew K. Gold (Ed.): *Debates in the Digital Humanities* (pp. 85–95). Minneapolis, London: University of Minnesota Press.
- Reh, S. (2014): Looking at Practices – Can We Discover Something New by Looking at Practices? Practice Theory and the History of Education. In: *Encounters on Education*, 15, 183–207.
- Rockwell, G. and S. Sinclair (2016): *Hermeneutica: Computer-Assisted Interpretation in the Humanities*. Cambridge, Mass.; London: The MIT Press.

Information Literacy Online

An Erasmus+ Project to Improve Students' Competencies

Stefan Dreisiebner

University of Graz, Austria
stefan.dreisiebner@uni-graz.at

Keywords: information literacy; MOOC; education projects

1 Motivation and background

Information Literacy (IL) is “the set of integrated abilities encompassing the reflective discovery of information, the understanding of how information is produced and valued, and the use of information in creating new knowledge and participating ethically in communities of learning” (American Library Association, 2016). IL as a social key competence is particularly essential in post-secondary education and research. According to many studies (e.g. Katz 2007; Rubinić et al., 2013), student’s information literacy levels are generally low. Analyses show the need for an attractive tutorial which can be developed on a broad basis of available material. Accordingly, in November 2016 the EU project ILO was started with the aim to develop, evaluate and disseminate a multilingual Massive Open Online Course (MOOC) for developing information literacy focusing on higher education students.

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 368–370.

2 Expected outcome

MOOCs are online courses with no entry barriers aiming at unlimited participation. The ILO project will at first concentrate on information literacy elements which are relevant for all subjects/disciplines. Examples for such general information literacy elements are Boolean operators, basic principles in knowledge organization, or basic knowledge of copyright law. IL also covers subject-specific elements, so the project will demonstrate the applicability of the generic information literacy MOOC to two exemplary disciplines: Business Administration and Psychology.

A major shortcoming of current IL courses is the lack of self-assessment components. Therefore, a central innovative approach of our MOOC will be the implementation of technology based assessment components which allow students to get feedback on their learning success. The MOOC content will be developed for six European cultural and language groups: English, German, Spanish, Catalan, Slovenian and Croatian. The multilingual approach will not only consider formal translation but also cultural-specific differences in the various realizations. The project will strongly endeavor to evaluate the MOOC in several phases and with different methods.

Acknowledgements

The author would like to thank the project team members Christian Schlögl (University of Graz), Thomas Mandl (University of Hildesheim), Alexander Botte (German Institute for Educational Research and Educational Information), Maja Žumer and Tanja Merčun Kariž (University of Ljubljana), Franjo Pehar, Ivanka Stričević and Mate Juric (University of Zadar), Lyn Robinson (City, University of London), and Ángel Borrego (University of Barcelona) for preparing the presented project concept.

The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

References

- American Library Association, ACRL Board (2016): Framework for Information Literacy for Higher Education. Chicago, ALA. <http://www.ala.org/acrl/standards/ilframework>
- Katz, R. I. (2007): Testing Information Literacy in Digital Environments: ETS's iSkills Assessment. In: *Information Technology and Libraries*, 26 (3), 3–12. <http://ejournals.bc.edu/ojs/index.php/ital/article/viewFile/3271/2884>
- Rubinić, D., I. Stričević and J. Mate (2013): Information Literacy Course – The Perception of Students and Professors: University of Zadar Case. In: Kurbanoglu Serap et al. (Eds.): *Worldwide Commonalities and Challenges in Information Literacy Research and Practice: European Conference on Information Literacy, ECIL 2013, Istanbul, Turkey, October 22–25, 2013, Revised Selected Papers* (pp. 528–534). Cham et al.: Springer International Publishing. <http://goo.gl/oa4F1C>

International Degree Mobility in Library and Information Science

Vera Hillebrand, Elke Greifeneder

Humboldt-Universität zu Berlin, Germany
vera.hillebrand@hu-berlin.de, greifeneder@ibi.hu-berlin.de

Abstract

This study explores patterns of the geographical mobility for researchers in Library and Information Science and shows that there are clear patterns towards the United States in particular, and more general to countries offering an English language education.

Keywords: degree mobility; education; LIS; lingua franca

1 Introduction and method

This study explores patterns of the geographical mobility for researchers in Library and Information Science (LIS) – a field where the major publication venues require proficient English language skills. The study is based on a dataset of 877 active LIS researchers who have been involved in the 2014 to 2016 iConferences. The dataset has been originally collected for Bogers & Greifeneder (2016) who studied the review process and the community of attendees and reviewers of the iConference. The dataset has been augmented by collecting manually additional demographic information on authors and reviewers: gender, if the PhD was completed, the doctoral granting university, the country and continent where the PhD was completed, the country and continent of the Master's degree and country and continent of origin.

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 371–373.

2 Results

The results of the study on geographical mobility patterns reveal alarming trends: 91.7% of all American researchers in the sample never left their home country to study or live abroad. International experience means getting to know more research communities, education styles and experiencing personal development through other languages and cultures. Of the 26 American researchers who moved away only six are today staying in a non-English-speaking country. On the other hand, researchers from Asia and Europe show a high rate of mobility towards North America. For Asia in particular women decided to leave the continent: 86.8% of all Asian females have left their country to receive an English degree in North America. More alarming, only 3.4% of all Asian researchers who have left their country at one point currently live in a non-English speaking country. Researchers born in Europe – if they have not left to stay in North America – are circulating on the continent itself with a preference for English speaking areas. Again 76.1% of all European researchers never left their home country. Middle East, Africa and Australia as minorities in the iSchools group are also losing active researchers to the United States. 41 of 43 researchers who left their country currently live in an English-speaking country and not one of all 43 researchers earned a degree in a non-English-speaking European country. The study on mobility in LIS shows very clearly that a majority of the next generation of LIS researchers receives an education in North America. 94.3% of all PhD students in the sample currently live in the US and may never return. One important pull factor seems to be the possibility of studying in English. If foreign students decide to come to Europe, they go to Ireland or the United Kingdom. Offering more degrees in English may be an important survival factor for European and Asian LIS schools.

3 Limitations

The results of this study are limited by the sample of active members of an international conference: in consequence the sample may contain more researchers who are willing to move to another country than if collected at a national conference in any of the European or Asian countries. On the other hand, the

sample includes those researchers who have come to play a role on the international level and as such it matters if they have a high or a low mobility index: having lived in another country or having an English speaking LIS education may be one of the core determinants of being a successful researcher in this field.

References

Bogers, T. and E. Greifeneder (2016): The iSchool Community: A Case Study of iConference Reviews. In: *iConference Proceedings*. <https://www.ideals.illinois.edu/handle/2142/89312>

Limits and Potentials of Grasping User Interaction in Controlled and Open World Environments

Contrasting PIAAC Large-Scale-Assessment Usage Data with Web Analytics

*Peter Böhm, Christoph Schindler,
Heiko Rölke, Marc Rittberger*

German Institute for International Educational Research, Germany
{boehm, schindler, roelke, rittberger}@dipf.de

Keywords: usage data; web analytics; log data; computer-based assessment; web portal; qualitative analysis

1 Motivation and background

Educational practices are becoming more and more digitalized and learning is realized in the open web as well as in controlled digital learning environments. This provides new opportunities not only for new interactive teaching methods, but also for tracking the interaction with learning materials for improving teaching and learning as well as for assessments (cf. Siemens & Long, 2011). On one hand, the field of Computer-based assessment (CBA) uses these new research capacities for improving competence assessments in controlled and validated test environments. On the other hand, in the open web, tasks are pursued in interactive situations and protocolled in logfiles. In this study, these two different situations of open and controlled environments are contrasted by analyzing the different fabrications and expressiveness' of

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 374–376.

data. The main research questions are: To what degree can open and closed world log data of learning tasks be compared? Where do they differ and what conclusions can be drawn regarding the grasping of user interaction and possible mutual optimizations of the log data gathering? By addressing these questions, hitherto disparate areas of research are linked.

2 Research design

PIAAC (Programme for the International Assessment of Adult Competencies) is used as a case study for a controlled environment of a large-scale assessment with CBA tools, while the German Education Server (GES)¹ serves as a case study for the open data environment. Here, like in the field of CBA, usage data is gathered but its expressiveness is limited because of the open, uncontrolled usage and usage data (Keil et al., 2015).

This study is methodologically based on an in-depth, qualitative comparison of the quantitative usage data of one PIAAC test item and a comparable usage scenario of the GES. This comparison includes the production conditions of the test items, the usage data, the fabrication and syntax of the data as well as the 'expressiveness' of metrics that can be calculated based on these data.

3 Expected outcome

By comparing these two different situations of data fabrications and the expressiveness of the data, we aim to analyze the limits and potentials of usage data from closed, highly controlled environments and open-world, uncontrolled websites. With open-world usage data, difficulties are expected in reliably recognizing visitors and the 'real' browsing behavior, while large quantities of log data are available. By contrast, CBA usage data is more controlled; in the case of PIAAC even extensive background information is

¹ See <http://www.eduserver.de/>.

available on the participants. The number of participants is usually smaller than the number of users of large websites, whereby the test is realized under ‘artificial’ test conditions. These first insights address an exchange on an epistemological level between both worlds (CBA and web analytics), advancing the examination of expressiveness of user interaction in different learning settings.

References

- Keil, S., P. Böhm and M. Rittberger (2015): Qualitative web analytics: New Insights into Navigation Analysis and User Behavior – A Case Study of the German Education Server. In: F. Pehar et al. (Eds.): *Re:inventing Information Science in the Networked Society*. Proceedings of the 14th International Symposium on Information Science (ISI 2015), Zadar, Croatia, 19th–21st May 2015 (pp. 252–263). Glückstadt: Hülsbusch.
- Siemens, G. and P. Long (2011): Penetrating the Fog: Analytics in Learning and Education. In: *EDUCAUSE review*, 46 (5), 30.

Lowering the Barriers to Entry into LIS Master Programs

The EINFOSE Project

Thomas Mandl¹, Tatjana Aparac Jelušić²

University of Hildesheim, Germany
mandl@uni-hildesheim.de

University of Osijek, Croatia
taparacjelusic@ffos.hr

Abstract

In most universities in Europe, the barriers to enter into a LIS master degree after earning a degree in another subject are typically high. The same is true for the international mobility of students enrolled in BA or MA LIS programs in Europe. The EINFOSE project (Information Science Education: Encouraging mobility and learning outcomes harmonization) aims at lowering this barriers by offering summer schools on core topics of information science and by providing open educational resources (OER) on these topics in the long term.

Keywords: LIS education; mobility barriers; curricula

1 Barriers

The European LIS programs are highly diverse and offer different approaches to dealing with complex digital information systems. Although they all prepare their students for handling information processes, the contents and

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 377–379.

structures of courses can be very heterogeneous despite some efforts for harmonization (Kajberg & Lørring, 2005; Johnson, 2013).

This leads to very different entry requirements for master programs. Students from other disciplines who wish to acquire a master degree in LIS based on the background of their first degree usually cannot easily enter a master program. There are several different approaches to handle such cases in Europe, e.g. the requirement to take BA courses before registering for the master or special courses to acquire the core competencies.

2 EINFOSE

The EINFOSE project aims at lowering this barrier and providing best practices for students from LIS and other BA programs. EINFOSE is a partnership of 8 universities funded under the Erasmus+ scheme. The partners are: University of Osijek, Croatia, Hoegskolan i Boras, Sweden, Karl-Franzens-Universität Graz, Austria, Univerza v Ljubljani, Slovenia, Università di Pisa, Italy, Universitat de Barcelona, Spain, Hacettepe Universitezi, Turkey and Stiftung Universität Hildesheim, Germany (<http://sokrat.ffos.hr/einfose>). EINFOSE runs until 2018.

Within EINFOSE, OER for core areas in LIS will be developed. They are intended to support students to close their knowledge gaps and to prepare better for MA programs. OER will be developed for the four following areas from the heterogeneous perspectives (in the languages of most partners): ‘Advances in Information Science’, ‘Research Methodology in Information Science’, ‘Principles of Information Seeking’ and ‘Retrieval and Evaluation of Information Services’.

Colleagues, who have good didactic approaches for teaching some specific topic, are invited to participate. Thus, the project promotes the exchange of good didactic practice. The new OERs will be tested and optimized at two Summer Schools. The first one takes place in August 2017 in Hildesheim. The Summer School covers the OER content and online communication between students and teachers including supervision of students’ assignments afterwards. The OER can also be used for further self-study. The creation of OERs should also lead to discussions about teaching and learning outcomes. This should also lead to a substantial reduction of barriers within Europe.

Based on the experiences, the project will assemble policy recommendations for universities and other stakeholders like accreditation agencies. These will be presented and discussed at an international conference organized by the EINFOSE team in September 2018 in Pisa.

References

- Johnson, I. (2013): The impact on education for librarianship and information studies of the Bologna process and related European Commission programmes – and some outstanding issues in Europe and beyond. In: *Education for Information*, 30, 63–92.
- Kajberg, L. and L. Lørring (Eds.) (2005): *European Curriculum Reflections on Library and Information Science Education*. Copenhagen: The Royal School of Library and Information Science. <https://www.asis.org/Bulletin/Dec-06/EuropeanLIS.pdf>

Open Encyclopedia System

Open Source Software for Open Access Encyclopedias

Nicolas Apostolopoulos, Christoph Schimmel, Ilker Egilmez

Center for Digital Systems, Freie Universität Berlin, Germany
{napo, cedis, iegilmez}@cedis.fu-berlin.de

Abstract

The main objective of the project “Open Encyclopedia System (OES)” is to develop a standardized web-based open source platform for building and maintaining scientific online encyclopedias, which will be publicly available, free of charge on the internet and accessible worldwide. The project consortium is carried out of four partners: Center for Digital Systems (CeDiS), Friedrich-Meinecke-Institut (FMI), Center for Modern Greece (CeMoG) all based at the Freie Universität Berlin, as well as the Bavarian State Library (BSB) in Munich.

Keywords: open access; e-publishing; online encyclopedia; versions; crowd sourcing; digital humanities; digital history

1 Designing open source software for open access encyclopedias

Between the years of 2011 and 2015, an English language online encyclopedia on the First World War “1914-1918-online” has been created, published and is freely accessible worldwide. The multi-perspective, open access encyclopedia is a collaborative project by the largest network of World War One

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 380–385.

researchers worldwide, spanning more than 50 countries. “1914-1918-online” features innovative navigation and search functions based on semantic wiki technology and will eventually contain around 1,600 articles. The well-known framework Semantic MediaWiki (SMW)¹ is used as an editorial system.

The follow-up project Open Encyclopedia System (OES) started in 2016 and will – by the end of the project in 2019 – offer a generic modular open source software system with which editors and users are able to build an online encyclopedia on their own. Our claim is to create a system for the users, in which collaboration is easy and web-based. The OES software should be customizable, flexible and adaptable to context applications and simple to operate. Ultimately, the OES software should set a standard in saving time and in being a sustainable framework.

After almost one year of analyzing the requirements for an Open Encyclopedia System and conceptualizing the software architecture and its components, we stand in between the design and the implementation phase. The core components will be ready for release by the end of spring 2017.

1.1 The necessity of an *Open Encyclopedia System*

Hitherto existing online encyclopedias however, prevalently stick to a linear and static content structure like a reversed image of print publications. There are to date no satisfying solutions that are able to meet the challenges of the digitized era nowadays: e.g. the dynamic sampling of electronic publications will need solutions for displaying various versions of articles alongside with reliable referencing and collation. Furthermore, the potential of crowd sourcing with reference to identifying updatable content, contributing (external) links or collaboratively extending bibliographies is not completely and consequently utilized as it is already being done in other web-based projects in the humanities.

In addition, most online encyclopedias are mere publication systems that do not make use of the internet’s possibilities of designing interactive working environments neither for editors nor for authors. Options for personalized settings, e.g. individualized search and navigation history, personal article lists or annotations, are usually missing. This may be considered as a huge deficit referring to online encyclopedias as these are mostly consisting of a

¹ https://www.semantic-mediawiki.org/wiki/Semantic_MediaWiki

huge quantity of data that is worked on long-term. Academic encyclopedias might benefit if they are able to connect high-quality, peer-reviewed knowledge with the advantages of publishing electronically in a persuasive way.

Until now, academic online encyclopedias are implemented as proprietary and singular versions. Most software components are tailored models for specific requirements of a use case. New projects aiming at publishing reference works online may therefore benefit from foremost developed concepts of digital publication (e.g. web presentation, search, navigation, etc.). When using aforesaid software for own projects, a client has two options: adopting specific project-related decisions or adapting the software in a costly way. Consequently, just using the same software would be either expensive or of less quality. Taking this into consideration, it is desirable to develop a standardized information architecture for the creation, publication and maintenance of open access online encyclopedias comparable to software solutions for other forms of publications like e.g. online journals (with the software OJS) or online monographs (with the software OMP). There is a necessity for an open and modular solution that can provide all functions that are adaptable and essential for designing a specific online reference work in a cost and time effective way.

1.2 Guiding principles of the software design

Our main concepts and visions are predominantly led by four principles – and are the key advantages over a wiki software extension (e.g. SMW) – in order to ensure the sustainability, adaptability, longevity and usability of such a software system: modularity, openness, integrity of data and interconnection. Our aim is to build a generic software system that is web-based and features a modular plug-in architecture. Furthermore, a key asset of the OES will be its availability as open source software. This will make the software easily modifiable and thus optimizable and might be kept up-to-date. Beyond that, one of the main challenges of the software development of OES will be to guarantee the interoperability between the software modules, components and interfaces.

2 Modules, publication and interfaces

The online platform of the Open Encyclopedia System will be built with a set of modules that are adaptable to specific application contexts. Each module will be (de-)selectable and composed by configurable components. The flexibility of modules and components ensures the adaptability to the requirements of specific needs of the digital publication of an encyclopedia. Additionally, the online platform will need to stay open source itself, therefore we will need to find a way of providing each technological component and feature without any financial obstacles. However, it should be possible for users to add proprietary software if it is advisable in a particular context.

The customizable modules of the OES platform will amongst others consist of components for workflow and project management, services for administering multimedia content, flexible assignment and management of user rights.

Individual and project-related workflows will be guaranteed with configurable tools and plug-ins. With providing various types of (open) peer review, automatic DOI registration and collation tools, academic standards will be held up on a high level. Numerous functionalities of community engagement and participation such as an integrated “call for papers”, blogs and comments that will be implemented within OES will strengthen the networks of scientists and authors and increase an encyclopedia’s impact.

2.1 Typical workflow of publication

The only existing typical workflow of publication so far is the one of the online encyclopedia “1914-1918-online. International Encyclopedia of the First World War”; in the course of the functional upgrade, the workflow will be modified. Below, a typical workflow scheme is sketched:

First of all, conceptualizing the articles and its structures will take place in a wiki environment, where only editors and the editorial office communicate. Applications to paper calls can be made with the integrated web form, while a list of open calls will be updated daily on the website of the encyclopedia. The authoring and reviewing process will be organized with the help of a project management tool that will be offered as a module and can be adapted to a context application. The process of editing and indexing will occur in the editorial system, which is part of the core components of the OES. Articles,

enrichment of metadata and versions of articles can be managed here. Each article will be connected to a specific bibliography that might be linked to Zotero² via a bidirectional interface. The publication platform will technically be separated from the other level(s) of data processing. In addition to the provision of encyclopedic articles, the publication system will include a personalized environment with reading lists, annotations, comments and shareable bibliographies.

2.2 Networking and interfaces

With reference to interconnection and networking, OES will provide interfaces to library catalogues such as OAI-PMH³, SWORD⁴, SRU⁵ et cetera in order to increase the visibility and impact of reference works created within the platform. Bibliographies that are created with the online reference tool Zotero may be edited via a bidirectional interface from and to a specific project.

In terms of availability, the Bavarian State Library (BSB) will offer professional services and strategies for the long-term preservation of digital objects documents tailored to particular encyclopedia projects.

3 Prospects

Currently, there are two use cases that will make use of the OES platform: a) “1914-1918-online. International Encyclopedia of the First World War”⁶, which was launched in October 2014 and is planned to be functionally upgraded with OES until 2018, b) the work-in-progress “Online Encyclopedia on German-Greek entanglements”, which should functionally be a scholarly publication platform for case study research and working papers on the history of cultural and scientific exchange between the German and Greek

2 <https://www.zotero.org/>

3 <https://www.openarchives.org/pmh/>

4 <http://swordapp.org/about/>

5 <http://www.loc.gov/standards/sru/>

6 <http://encyclopedia.1914-1918-online.net/home.html>

spaces from the 18th century to the present day (and is mainly planned by the Center for Modern Greece).

References

Ute Daniel et al. (Eds.) (2014): 1914-1918-online. International Encyclopedia of the First World War, issued by Freie Universität Berlin.

VICI

Visualisation of Collaboration in Social Enterprise Software Systems

Fabian Odoni, Wolfgang Semar, Elena Mastrandrea

Swiss Institute for Information Research (HTW Chur), Switzerland
{fabian.odoni, wolfgang.semar, elena.mastrandrea}@htwchur.ch

Abstract

This poster describes the relationship visualisation framework VICI. VICI was developed to extract co-author relationships found on an Atlassian Confluence installation and visualise them as network graphs. Therefore wiki-entries are extracted directly from the Confluence database using sql-queries. This data is used to calculate co-authorship bigrams and build a relationship graph. The graph is then visualised as force-directed network graph using D3.js in a web app. Node size and colour as well as path width and colour are used to encode relationship degrees and author activity. Additional relationship visualisations will be developed in the future.

Keywords: information retrieval; data visualisation; knowledge management; software development

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 386–388.

1 Introduction

In cooperation with a project partner, the Coltero project¹ aims to extract key performance indicators from the social collaborative knowledge management software Atlassian Confluence² and use them to develop a knowledge sharing incentive system.

With VICI, a first step was made to extract and visualise co-authorship relations from wiki entries in Confluence. The visualisation was already used to monitor the participation rate of the employees and thus estimating the degree of acceptance of the social enterprise software.

2 Architecture and methods

VICI is composed of a Python³ based backend system that extracts the data and calculates the relationships needed for the visualisations, and a web based frontend that displays the relationship graph and allows some user interaction with the graph as well as the backend system.

2.1 Data extraction

At the moment VICI focuses on the wiki-entries written in the Confluence wikis. The data extraction has to take into account that Confluence does not only store the latest version of a wiki-entry but also all its revisions. These revisions are linked by id numbers.

2.2 Calculating the relationship graph

Since every author of a wiki-entry is a co-author of every other author of the same wiki entry, a combination $([A, B, C] \rightarrow [[A, B], [A, C], [B, C]])$ of all

1 <http://www.htwchur.ch/digital-science/forschung-und-dienstleistung/institut-sii/projekt-uebersicht/coltero-collaboration-and-enterprise-knowledge-visualisation.html>

2 <https://atlassian.com/software/confluence>

3 <https://www.python.org>

the authors per wiki-entry is calculated. These bigrams are then entered with additional data like edge weight into a NetworkX⁴ graph.

2.3 Visualising the graph

The frontend graph visualisation uses the D3.js⁵ library to build a force-directed graph. Authors are displayed as nodes, and the relationship between them are shown as links between these nodes.

2.4 Web framework

In order to run the whole pipeline, a small web application was built using the Flask⁶ microframework. This application fetches and calculates the needed data, provides a web interface to display the visualisation and offer the option to switch between different databases.

3 Future development

Future development of VICI will include graphic and algorithmic improvements to the existing framework as well as the development of algorithms to display new relationships.

4 <https://networkx.github.io>

5 <https://d3js.org>

6 <http://flask.pocoo.org>

Research Data at the Faculty of Arts and Humanities of the University of Cologne

Simone Kronenwett

Data Center for the Humanities, University of Cologne, Germany
simone.kronenwett@uni-koeln.de

Keywords: research data; Arts and Humanities; RDM

In the course of the digitization of information, research data management (RDM) has become one of the most important new areas of research. Going one step further, Berman states that research data will drive the next generation of innovation and therefore, the development of effective research data infrastructure will be essential to enable data access and use (Ray, 2014: 438).

But according to estimates by the German Research Foundation (DFG), up to 90% of the digital generated research data and results end up getting lost (Winkler-Nees, 2011: 5) or “disappear in the drawer” (Kramer, 2014) shortly after the completion of research projects and are therefore not available for further use and reuse. To tackle this emergency, the German Council for Scientific Information Infrastructures (RfII) has made a series of recommendations for the future management of research data in its recently published position paper called “Performance by Diversity” (RfII, 2016). The RfII was tasked by Germany’s Joint Science Conference (GWK) with the formulation of broad-based recommendations for the scientific system in Germany as a whole. To that effect, universities have to prepare themselves to provide their scientists with the necessary infrastructures and services regarding RDM. The managing boards of the German universities organized in the German Rector’s Conference (HRK) have already identified this as a key task and published recommendations of how to develop an institutional RDM

In: M. Gäde/V. Trkulja/V. Petras (Eds.): Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, 13th–15th March 2017. Glückstadt: Verlag Werner Hülsbusch, pp. 389–391.

policy in each university (HRK, 2014, 2015). One of the recommendations of actions includes the conduction of an online survey addressing researchers' data and demands. The present poster acts on that suggestion by carrying out an online survey on RDM at the Faculty of Arts and Humanities of the University of Cologne. The survey was conducted in 2016 by the Data Center for the Humanities (DCH) in cooperation with the Office of the Dean of the Faculty of Arts and Humanities as well as the University and City Library of Cologne (DCH, 2016).

Because surveys in human-computer interaction (HCI) research can be useful to gather information about people's habits, interaction with technology, or behavior as well as to get feedback on people's experience with a product, service, or application and to collect people's attitude and perceptions towards an application in the context of usage, the methodical approach of the survey follows the six stages of "Survey Research in HCI" published by Müller, Sedley and Ferrall-Nunge (2014). In addition, the survey is based on the relevant articles published in the handbook "Methods of Library and Information Science" by Umlauf, Fühles-Ubach and Seadle (2013).

The objective in the compilation of the questionnaire was to answer the following questions:

1. What research data are available?
2. What is the need for research data?
3. What support do the members of the Faculty of Arts and Humanities want from the DCH?

The poster will focus the main results of the survey and show potential conclusions.

References

- Averkamp, S., X. Gu and B. Rogers (2014): Data Management at the University of Iowa: A University Libraries Report on Campus Research Data Needs. University of Iowa. http://ir.uiowa.edu/lib_pubs/153/ <1.11.2016>
- Bauer, B., A. Ferus, J. Gorraiz et al. (2015): Researchers and their Data. Results of an Austrian Survey – Report 2015. e-Infrastructures Austria, Vienna. <http://e-infrastructures.at/das-projekt/deliverables> <1.11.2016>

Research Data at the Faculty of Arts and Humanities of the University ... 391

Data Center for the Humanities (DCH) (2016): Research Data Survey 2016.
<http://dch.phil-fak.uni-koeln.de/umfrage-2016.html> <1.11.2016>

Hochschulrektorenkonferenz (HRK) (2014): Management von Forschungsdaten als strategische Aufgabe der Hochschulleitungen. Empfehlung der 16. HRK-Mitgliederversammlung am 13. Mai 2014 in Frankfurt am Main. https://www.hrk.de/uploads/tx_szconvention/HRK_Empfehlung_Forschungsdaten_13052014_01.pdf <1.11.2016>

Weitere Bände der „Schriften zur Informationswissenschaft“

Terje Tüür-Fröhlich: The Non-trivial Effects of Trivial Errors in Scientific Communication and Evaluation (Bd. 69; 2016) 164 S., 978-3-86488-104-6, 24,80 €

Ben Heuwing: Usability-Ergebnisse als Wissensressource in Organisationen (Bd. 68; 2015) 460 S., ISBN 978-3-86488-084-1, 35,80 €

Ulrich Herb: Open Science in der Soziologie Eine interdisziplinäre Bestandsaufnahme zur offenen Wissenschaft und eine Untersuchung ihrer Verbreitung in der Soziologie (Bd. 67; 2015) 500 S., ISBN 978-3-86488-083-4, 36,80 €

Franjo Pehar, Christian Schlögl, Christian Wolff (Eds.): Re:inventing Information Science in the Networked Society Beiträge ISI 2015 (Bd. 66; 2015) 596 S., ISBN 978-3-86488-081-0, 38,50 €

Anke Reinhold: Forschungsdaten in der videobasierten Unterrichtsforschung Benutzerzentrierte Modellierung und Evaluierung einer Domänen-Ontologie (Bd. 65; 2015) 280 S., ISBN 978-3-86488-061-2, 29,90 €

Norbert Henrichs: Menschsein im Informationszeitalter Informationswissenschaft mit Leidenschaft und missionarischem Eifer [Pioniere der Informationswissenschaft | Norbert Henrichs] (Bd. 64; 2014) 404 S., 978-3-86488-061-2, 33,90 €

Hans-Christoph Hobohm (Hg.): Informationswissenschaft zwischen virtueller Infrastruktur und materiellen Lebenswelten Beiträge ISI 2013 (Bd. 63; 2013) 502 S., ISBN 978-3-86488-035-3, 34,50 €

Saskia-Janina Untiet-Kepp: Adaptives Feedback zur Unterstützung in kollaborativen Lernumgebungen (Bd. 62; 2012) 322 S., 978-3-86488-023-0, 30,90 €

Carola Carstens: Ontology Based Query Expansion Retrieval Support for the Domain of Educational Research (Bd. 60; 2012) 396 S., 978-3-86488-011-7, 34,90 €

Matthias Görtz: Social Software as a Source of Information in the Workplace Modeling Information Seeking Behavior of Young Professionals in Management Consulting (Bd. 59; 2011) 314 S., ISBN 978-3-86488-006-3, 31,90 €

Joachim Griesbaum, Thomas Mandl, Christa Womser-Hacker (Hg.): Information und Wissen: global, sozial und frei? Beiträge ISI 2011 (Bd. 58; 2011) 536 S., ISBN 978-3-940317-91-9, 36,50 €

Alexander Warta: Kollaboratives Wissensmanagement in Unternehmen Indikatoren für Erfolg und Akzeptanz am Beispiel von Wikis (Bd. 57; 2011) 304 S., ISBN 978-3-940317-90-2, 30,90 €

Andreas Ratzka: Patternbasiertes User Interface Design für multimodale Interaktion (Bd. 55; 2010) 400 S., ISBN 978-3-940317-62-9, ~~33,90 €~~ UVP: 10,80 €

Thomas Memmel: User Interface Specification for Interactive Software Systems (Bd. 54; 2009) 362 S., ISBN 978-3-940317-53-7, ~~33,90 €~~ UVP: 11,80 €

Monika Maßun: Collaborative Information Management in Enterprises (Bd. 52; 2009) 222 S., ISBN 978-3-940317-49-0, ~~28,90 €~~ UVP: 9,80 €

Susanne Mühlbacher: Information Literacy in Enterprises (Bd. 51; 2009) 342 S., ISBN 978-3-940317-45-2, ~~32,90 €~~ UVP: 11,80 €

Rainer Kuhlen (Hg.): Information: Droge, Ware oder Commons? Wertschöpfungs- und Transformationsprozesse auf den Informationsmärkten Beiträge ISI 2009 (Bd. 50; 2009) 578 S., ISBN 978-3-940317-43-8, 38,90 €

Markus Heckner: Tagging, Rating, Posting Studying Forms of User Contribution for Web-based Information Management and Information Retrieval (Bd. 49; 2009) 244 S., ISBN 978-3-940317-39-1, 27,90 €

Everything Changes, Everything Stays the Same? Understanding Information Spaces

The 15th International Symposium of Information Science, a biannual conference for scholars, researchers, professionals and students in information science, takes place at Humboldt-Universität zu Berlin from 13 to 15 March 2017. This year's theme of "Everything Changes, Everything Stays the Same? Understanding Information Spaces" puts the dynamic of the field of information and its applications in the center.

With the ongoing digitization and virtualization of goods, services and living environments, information science reflects on the potential changes within the information society. While some parts of the community proclaim a revolutionary shift not only in the way we approach information and information systems, but also in the way society constitutes itself, others state that while the applications and interfaces adapt to advances in information technology, the underlying principles for human interactions with information remain the same.

In the ISI 2017 proceedings, contributions on the development of innovative information spaces and services, on analyses of human-computer interactions in physical or virtual information spaces and on the role of information science and its research endeavors are included.

www.vwh-verlag.de

Verlag Werner Hülsbusch
Fachverlag für Medien-
technik und -wirtschaft

ISBN: 978-3-86488-117-6



ISSN: 0938-8710
ISBN: 978-3-86488-117-6

33,90 € (D)
34,85 € (A)
38,90 CHF