Wolfgang Aigner
Kerstin Blumenstein
Michael Iber
Grischa Schmiedl
Matthias Zeppelzauer
(Eds.)

FMT 2016

9th Forum Media Technology 2016

2nd All Around Audio Symposium 2016

November 23-24, 2016
St. Pölten, Austria

Proceedings

Editors' addresses:
St. Pölten University of Applied Sciences
Institute of Creative\Media/Technologies
Matthias Corvinus-Straße 15
3100 St. Pölten
Austria

{wolfgang.aigner| grischa.schmiedl | kerstin.blumenstein | matthias.zeppelzauer | michael.iber}@fhstp.ac.at

Preface

Audiovisual, interactive and mobile media interweave with our everyday life. In almost all our areas of life, they change our behavior and affect our perception, thinking and feeling. The 'Forum Media Technology (FMT)' is a scientific event that aims for an intensive dialogue of experts working at agencies, studios and production companies with students, lecturers, researchers and developers in the field of digital media.

In its 9th edition, the Forum Media Technology was held on November 23-24, 2016 at St. Pölten University of Applied Sciences, Austria. In conjunction with the main conference, the 2nd edition of the trans-disciplinary symposium 'All Around Audio (AAA)' as well as the 1st edition of the new format 'Graduate Consortium' were part of the two day program.

All Around Audio Symposium

Although, audio has not abandoned its status as a standalone discipline, its trans-disciplinary participation at the conception and design of products and environments has become more and more essential. In this sense, All Around Audio not only addresses specialists in the audio domain but particularly encourages researchers and designers from other fields to participate in the symposium. In its second edition, 16 international speakers were invited for talks on a wide spectrum of topics reaching from media economy, audio technology, auditory display to music and media arts.

Graduate Consortium

The FMT 2016 Graduate Consortium session was intended to provide an opportunity for graduate and PhD students to explore, discuss and develop their research topic in an interdisciplinary workshop, under the guidance of a panel of distinguished researchers.

Forum Media Technology Conference Track

Submissions for the FMT conference track were accepted in two categories of full and short papers. All submitted papers underwent a double-blind review process where each paper was reviewed by at least three members of the international program committee (IPC) of the conference. Based on the written reviews, final decisions were made by the paper chairs and 12 papers (8 full + 4 short) were accepted for presentation at the conference (acceptance rate: 48%). The accepted papers can be clustered into four areas: data modeling & analysis, usability and mobile applications, information visualization, as well as digital media experience.

For the first time, also one best paper and two honorable mention awards were given to the top three contributions of the FMT. The award is based on the scientific excellence of the paper as well as the presentation performance of the authors at the conference. The selection process for the best paper award followed a three-step process. First, IPC members indicated whether a paper should be considered for the award. Based on that, the paper chairs decided on a short list of three nominated papers for the award and the selected papers' authors were informed about their nomination. The final decision was made by an anonymous best paper committee consisting of three members who took into account both the paper as well as the presentation to make their final decision.

November 2016

Wolfgang Aigner
Grischa Schmiedl
Kerstin Blumenstein
Matthias Zeppelzauer
Michael Iber

Best Paper Award Sponsor:
WISSENSCHAFT · FORSCHUNG NIEDERÖSTERREICH

Organizing Committee

Conference Chairs

Wolfgang Aigner, St. Pölten UAS, Austria
Grischa Schmiedl, St. Pölten UAS, Austria

Paper Chairs

Kerstin Blumenstein, St. Pölten UAS, Austria
Matthias Zeppelzauer, St. Pölten UAS, Austria

Graduate Consortium Chair

Grischa Schmiedl, St. Pölten UAS, Austria

AAA Symposium Chair

Michael Iber, St. Pölten UAS, Austria

Local Organizers

Rosmarie Tomasch, St. Pölten UAS, Austria
Katharina Schloder, St. Pölten UAS, Austria
Ulrike Wieländer, St. Pölten UAS, Austria

Best Paper Committee

Andreas Gebesmair, St. Pölten UAS, Austria
Cornelius Pöpel, Ansbach UAS, Germany
Michael Zeiller, UAS Burgenland, Austria

Program Committee

Bilal Alsallakh, Bosch Research, San Francisco, USA
Arnold Baca, Vienna University, Austria
Benjamin Bach, Monash University, Australia
Alessio Bertone, Dresden University of Technology, Germany
Jakob Doppler, St. Pölten UAS, Austria
Astrid Ebner-Zarl, St. Pölten UAS, Austria
Stefan Emrich, Drahtwarenhandlung, Austria
Thomas Felberbauer, St. Pölten UAS, Austria
Mario Heller, Vienna University, Austria
Andreas Gebesmair, St. Pölten UAS, Austria
Pascal Goffin, Inria, France
Nicole Gonser, FHWien UAS of WKW, Austria
Thomas Grill, Austrian Research Institute for Artificial Intelligence, Austria
Johanna Grüblbauer, St. Pölten UAS, Austria
Robert Gutounig, FH Joanneum UAS, Austria
Svenja Hagenhoff, Friedrich-Alexander University Erlangen-Nürnberg, Germany
Isabel John, UAS Würzburg-Schweinfurt, Germany
Andreas Kerren, Linnaeus University, Sweden
Simone Kriglstein, Vienna University, Austria
Georg Kuntner, St. Pölten UAS, Austria
Tatiana von Landesberger, TU Darmstadt, Germany
Robert Laramee, Swansea University, UK
Christian Märtin, University of Applied Sciences Augsburg, Germany
Dalibor Mitrović, mediamid digital services GmbH, Austria
Paul Modler, Karlsruhe University of Arts and Design, Germany
Thomas Moser, St. Pölten UAS, Austria
Paul Murschetz, University of Klagenfurt, Austria
Christoph Musik, St. Pölten UAS, Austria
Tassilo Pellegrini, St. Pölten UAS, Austria
Peter Plessas, Institute of Electronic Music and Acoustics Graz, Austria
Cornelius Pöpel, Ansbach UAS, Germany
Helmut Ritschl, FH Joanneum UAS, Austria
Paul Rosenthal, Chemnitz University of Technology, Germany
Raphael Schneeberger, PocketScience GmbH, Austria
Michael Sedlmair, University of Vienna, Austria
Rosa von Suess, St. Pölten UAS, Austria
Christian Tominski, University of Rostock, Germany
Manuela Waldner, Vienna University of Technology, Austria
Michael Zeiller, UAS Burgenland, Austria

Contents

Keynote & Capstone

Keynote: Visual-Interactive Data Analysis – Obtaining Insight into Complex Data Using Interaction, Visualization, and Data Mining

Tobias Schreck, TU Graz, Austria

Abstract

Advances in data acquisition and storage technology lead to the creation of increasingly large, complex data sets across application domains as diverse as science, engineering, business, social media, or team sports analysis. Important user tasks for leveraging large, complex data sets include finding relevant information, exploring for patterns and insights, and re-using of data for authoring purposes. Novel methods in visual-interactive data analysis allow to tightly integrate knowledge of domain analysts with automatic data analysis methods, offering solutions for complex analysis problems. We discuss visual-interactive data analysis techniques from our work that support search and analysis in a variety of different data types and novel application scenarios. These include approaches for example- and sketch-based search in multidimensional data sets, exploration of patterns in social media data, and visual analysis of soccer match data. We also touch on novel methods to support the analysis and restoration tasks in virtual archaeology. We conclude the talk by discussing research challenges in the area.

Biographie

Tobias Schreck holds a Professor position with the Institute for Computer Graphics and Knowledge Visualization at Graz University of Technology, Austria. Between 2011 and 2015, he was an Assistant Professor with the Data Analysis and Visualization Group at University of Konstanz, Germany. Between 2007 and 2011 he was a Postdoc researcher and head of a junior research group on Visual Search and Analysis with Technische Universität Darmstadt, Germany. Tobias Schreck obtained a PhD in Computer Science in 2006, and a Master of Science degree in Information Engineering in 2002, both from the University of Konstanz. He works in the areas of Visual Analytics, 3D Object Retrieval, and Digital Libraries. His research interests include visual search and analysis in time-oriented, high-dimensional and 3D object data, with applications in data analysis, multimedia retrieval and cultural heritage. He has served as co-chair for Posters, Workshops and Panels for IEEE VIS, as well as a co-organizer for the EG Workshop on 3D Object Retrieval in the past.

Capstone: How to bring people to new worlds?

Andreas Jakl, Tieto, Austria

Abstract

Each tool modifies, extends or improves a tiny aspect of our daily life. That's the goal of every mobile app – Augmented Reality and Wearables are just the most recent hypes. However: what are the possibilities to inject actually useful information into the world of the user? The planning of remarkable apps that achieve exactly that does already start with the user experience and combines it in a novel way with other immersive technologies like location based services. We will take a look at the recipes for success and inspiring apps that managed to build this bridge!

Biographie

Andreas Jakl is working with mobile apps since 2004. As Mobility Expert at Tieto, he brings the business partner's vision of mobile strategies to life. Together with the mobility.builders Community, he is organizing the Mobile Developer After-Work events, where expert developers share their knowledge. As Microsoft MVP (Most Valuable Professional) for Windows Development, he publishes open source libraries for immersive technologies like Bluetooth Beacons and NFC.

Session 1: Data Modeling and Analysis

Adoption of Technical Reporting Standards Among Austrian listed Companies – The Case of XBRL

Monika Kovarova-Simecek
UAS St. Poelten
Matthias Coprvinus Str. 15
3100 St. Poelten, Austria
Monika.Kovarova-Simecek@fhstp.ac.at

Tassilo Pellegrini
UAS St. Poelten
Matthias Coprvinus Str. 15
3100 St. Poelten, Austria
Tassilo.Pellegrini@fhstp.ac.at

Abstract

The requirements towards financial reporting (FR) have considerably changed within the last 15 years. Stakeholders demand not only accurate and reliable information in shorter intervals, but also customized reports meeting their information needs. Thus, companies need to develop strategies to cope with the new affordances of professional investor relations and stakeholder management. We conducted a survey among publicly listed Austrian firms, investigating whether they perceive a need to develop new reporting practices and if they have already started to deal with new sorts of reporting standards, especially XBRL. The survey examined the state of the art in XBRL diffusion and adoption among Austrian companies analysing supporting and inhibiting factors for its application and rejection. The results of the survey indicate a great awareness for the need of target-group oriented financial reporting and a high relevance of technical reporting standards in the future. However, Austrian firms show poor preparedness for the new technological requirements. It's probable that initiatives are needed to stimulate the adoption of the new technological standards and pave the way towards a next generation reporting.

1 INTRODUCTION

With the increasing proliferation of the Internet as a universal medium for information exchange and presentation the affordances of financial reporting (FR) of publicly listed companies have changed. As various stakeholders along the information value chain demand more information in shorter intervals [15], companies have to develop new reporting strategies that transcend the limitations of static, paper-based reporting and harness the capabilities of digital publishing media. For over a decade companies have used a variety of electronic publication formats to provide financial data to the public. Formats like PDF and HTML have gained a broad acceptance among the investor relations community and are being used widely for documentation and communication purposes. But as stated by Rodriguez [17], "(...) investors are explicitly given prominence on the website and although ample investor relation information is provided, the attention to investor relations is not exclusive, and there are other stakeholders featured on the companies' websites", like consumers, employees or regulatory agencies. All these stakeholders have differing information needs, and it is difficult to meet these needs by one standardized financial report. Hence, conventional formats go hand in hand with certain deficiencies when it comes to the customization of reports for specific target groups and the flexible reusability of financial data contained in these publications. In short, conventional technologies limit the scale and scope of reporting innovations, making it difficult to react to the changing affordances of the financial reporting environment.

Over recent years, various business reporting standards have been developed that among other things address the reuse of financial data. The most comprehensive and mature format is XBRL, the eXtensible Business Reporting Language, an expressive XML-vocabulary optimized to represent financial data at a highly granular level. XBRL separates the presentation layer from the data contained in it, and thus increases the usability of financial data for purposes such as reporting, analytics and targeted contextualisation. Dunne et al. [4] argue that: "Documents rendered by XBRL are digitally-enabled so that it is easier for stakeholders to extract information directly into spreadsheets, or any other XBRL-enabled software, without the need to re-key data thus providing significant improvements in information flows and enhancing inter-company comparability." Accordingly, XBRL is perceived to be a promising standard that meets the requirements of new reporting routines and also challenges existing (de-facto) standards in the domain of financial reporting [4;9;18].

This paper contributes to the increasing number of works investigating the diffusion of XBRL as an enabling technology for new reporting routines and practices.

In: W. Aigner, G. Schmiedl, K. Blumenstein, M. Zeppelzauer (eds.): Proceedings of the 9th Forum Media Technology 2016, St. Pölten, Austria, 24-11-2016, published at http://ceur-ws.org

Several country-specific studies have already been provided (for details see sec. 2), but no attention has so far been paid to the adoption of XBRL in Austria.

To close this gap, we conducted a survey among publicly listed Austrian firms whether they perceive a need to develop new reporting practices and if they have already started to adjust to the new circumstances. Aside these general insights, the survey investigated the state of the art in XBRL adoption among Austrian companies, analysing supporting and inhibiting factors for its application and rejection.

The paper is structured as follows: Chapter 2 gives a brief introduction into XBRL, explaining its evolution and core features. Chapter 3 discusses related work dealing with the diffusion and adoption of XBRL and associated institutional setups. In chapter 4 the authors explain the survey's methodology and present the survey results. Chapter 5 provides a discussion and conclusion.

2 EXTENDED BUSINESS REPORTING LANGUAGE – DESCRIPTION AND EVOLUTION

Since 1999 the US based company XBRL International Inc. has been standardising XBRL currently providing it to the public under version 2.1. XBRL is a scripting language based on XML "intended for modelling, exchanging and automatically processing business and financial information" [7]. XBRL allows representing financial metadata in a standardized, machine-processable form by linking reporting facts to standard financial taxonomies (such as IFRS[1] and US-GAAP[2]) and extend these statements with individual metadata according to a company's specific reporting needs [2]. Thus, XBRL allows maximum flexibility in the contextualization and reuse of financial data for various reporting purposes [8]. XBRL should be considered as a specific reporting extension to general purpose electronic business languages like EDIFACT (Electronic Data Interchange for Administration, Commerce and Transport) or ebXML (Electronic Business using eXtensible Markup Language) whose main purpose is to represent and align processes between supply chain partners in a standardized way.

Recently, new methodologies have been introduced to further extend the expressivity of XBRL by enriching it with other standardized vocabularies and data sources. This so called Linked Data approach [16;7] is a profound technological leap in the customization of financial reports according to the specific needs of various target groups. As noted by Guillox et al. [9], "(…) the extensibility offers a role back to the human in the process of instituting regulatory procedures and filing submissions". Investors, suppliers, employees, customers, regulators, financial analysts, researchers might receive comprehensive, yet customized financial data without selecting the data needed from static documents provided in paper or PDF. In addition, this new approach would avoid that firms can filter financial information provided to stakeholders and e.g. present less favourable information in footnotes of financial reports which are not as strongly received as the main body of financial reports [12;18]. XBRL promises to improve the transparency and accuracy of financial reporting and allows a higher protection for financial data users. If companies manage to reach stakeholders in an intelligible way, they gain their trust and could enhance the company value [14].

Despite several efforts to establish XBRL as an electronic reporting standard, its broad adoption is still in its infancy and its impact is still subject to debate. In the following sections we provide a comparison between the United States of America and Europe according to similarities and differences in the adoption of XBRL. According to Kernan [11], "XBRL is evolving everywhere, but unevenly, driven by various stakeholders such as governments, stock exchanges, banks and other industry sectors".

2.1. XBRL Diffusion in the US

In the United States of America the Securities and Exchange Commission (SEC) has started in 2009 to use XBRL as mandatory reporting standard for electronic records, thus stimulating the steady uptake of XBRL among US publicly listed companies [19]. Prior to this in 2008 the Federal Deposit Insurance Corporation (FDIC), a public agency assessing risks in the nation's financial system, started to collect XBRL records from over 8000 banks on a quarterly basis [11]. Since then, numerous studies investigated the impact of XBRL diffusion among the US financial industry. Some of the latest results are presented below.

Baldwin & Trinkle [1] interviewed a Delphi panel on the potential impacts of XBRL on the financial industry. They conclude that "XBRL is very likely to impact corporations, financial reporting, users of financial reports and auditing. The most likely impacts of XBRL include: increased accessibility of financial reports, easier regulatory compliance, enhanced availability of financial reports, facilitation of continuous reporting, and improved efficiency in investment and business decision making."

Sinnet [20] conducted a survey among 442 US companies and concludes that XBRL literacy among US companies is rising. According to his findings, "companies have reduced the amount of outsourcing services used to create their XBRL filings, and they expect to further reduce outsourcing over the coming year. Significantly, over half of large accelerated filers do not expect to use XBRL professional services for their next annual filing. This trend suggests that larger filers continue to become confident that they can be self-sufficient with the preparation and review of their XBRL reports."

By analyzing the impact of XBRL on analyst forecast behaviour Liu et al. [13] found "a significant positive association between mandatory XBRL adoption and both

[1] See also http://www.ifrs.org, accessed 2016-10-10
[2] See also http://usgaap.pro/, accessed 2016-10-10

analyst following and forecast accuracy." According to the authors "the findings not only support the SEC's requirement of detailed tagging of footnotes but also show that the benefits of adopting XBRL are realized regardless of errors found and concerns raised at the early stage of adoption" (ibid.).

Interestingly, Dhole et al. [3] come to a somewhat contradictory conclusion. Their survey results conducted among US XBRL filings indicates that the existing adoption of XBRL among US companies lead to a decline of financial statement comparability, also due to the company-specific extension taxonomies. Additionally, they found that selling, general and administrative expense comparability declined after the mandate, while depreciation comparability did not change.

2.2. XBRL Diffusion in Europe

In Europe the circumstances for the diffusion of XBRL differ profoundly as compared to the US. It is characterised by a nationally fragmented, regulatory landscape, making it difficult to establish a common reporting standard throughout the European Union. In a workshop conducted in 2011 by the financial service provider ICAEW and the University of Birmingham the organizers came to the conclusion that "[...] there are significant barriers to a pan-European adoption of XBRL for company reporting in the style of the U.S. SEC's mandatory requirement. The democratic right of member states to determine their own filing arrangements (through Officially Appointed Mechanisms) is both a vital core principle of the EU's operating practices and yet a barrier to a timely and effective response to the challenge of pan European security market supervision, in which XBRL could play a role. It is also important to take into account that different regulator implementations have different goals, which must be well defined to determine precisely what is to be made mandatory" [10].

To overcome these obstacles various initiatives have been launched at the national and international level to promote the adoption of XBRL. At the international level the European Committee of Central Balance-Sheet Data Offices (ECCBSO) has established the ERICA working group to monitor the usefulness of XBRL as a tool to reduce the reporting burden for IFRS. The group is chaired by the Banco de España and comprised of the following members: Banco de Portugal, Banque de France, Banque Nationale de Belgique - Nationale Bank van België, Cerved Group spa - Centrale dei Bilanci, Banca d'Italia, Deutsche Bundesbank, Oesterreichische Nationalbank, Bank of Greece and the European Central Bank. In an activity report from 2010 they come to the conclusion that "[...] the European commitment to XBRL has meant the creation of the XBRL Europe entity, with the aim of coordinating the efforts of the different European XBRL jurisdictions. Finally, some Central Balance Sheet Data offices belonging to the Committee have developed and are continuing to play a key role in the diffusion of XBRL as a new tool for dissemination of financial information in their countries; [...]" [5].

In 2007 Rodriguez et al. [17] conducted a study on financial reporting strategies among Spanish regional governments. Back then, none of the surveyed 13 governmental bodies used XBRL, XML or XLS for the disclosure of financial information. The authors come to the conclusion that "new technologies such as the Internet are not relevant for Spanish regional governments as a means of disclosing their financial information among the different users" (ibid., p. 163). Since then various initiatives originating from the Bank of Spain in cooperation with the Ministry of Industry, Tourism and Commerce have taken place whose aim it was to stimulate the adoption of XBRL among the public and the private sector. According to Escobar-Rodriguez & Gago Rodriguez [6] "the use of the standard is spreading to all areas. In the public sector, taxonomy for the rendering of accounts by the Local Entities of the Ministry of Economy and Finance has been developed, on the initiative of the General Inspectorate of the Administration of the State, the Ministry of Economy and Finance, and the General Directorate of Financial Coordination with the Autonomous Communities and with Local Entities. In the private sector, the taxonomies of the Institute of Accounting and Auditing of Accounts of the Ministry of Economy and Finance (ICAC) and of the National Commission of the Securities Market (CNMV) are significant."

Guilloux et al. [9] investigate the contestation of two technical reporting standards - EDIFACT and XBRL - among French government agencies for purposes of collecting business data for regulatory purposes. By conducting an actor-network-analysis the authors illustrated the institutional diffusion of XBRL as an informal competitor to the official EDIFACT standard. According to their findings "[s]ome proponents originally believed that companies would voluntarily adopt XBRL to enhance information for investors, but it came apparent that only regulators had a clear business case for adoption and businesses would not volunteer to be accountable" (ibid., 269). They conclude that "the newness of XBRL's technology just as regulators need to respond to an economic crisis and its [XBRL] adoption by French regulators not using EDIFACT create an opportunity for the challenger to make significant network gains over the long term" (ibid., p. 257).

For the UK Dunne et al. [4] collected 1733 questionnaires from business accountants, tax practitioners, auditors and financial professionals. They come to the conclusion that "awareness of XBRL, and second generation reporting more generally, resides in key champions but there is little diffusion outside this narrow set of stakeholders. Regulatory engagement seems to be the only impetus for diffusion and better channels of communication within stakeholder networks, such as between regulators, preparers, users and the XBRL community are needed" (ibid., p. 167)

This brief overview of the XBRL diffusion in the US and Europe outlines a twofold scenario. On the one side, we see various governmental initiatives that aim at stimulating the adoption of XBRL as technical reporting standard, on the other side awareness about XBRL exists, but the voluntary uptake of XBRL by companies and their stakeholders is lagging despite the multiple benefits of the standard in fulfilling the requirements of a "second generation reporting" [4]. ICAEW [10] conclude that "[t]agging business data using XBRL is part of the larger movement to create a semantic web to free data for exchange and automated re-use. It has made significant progress, but faces important institutional and infrastructure challenges in becoming ubiquitous in business reporting settings in Europe."

3. ADOPTION OF XBRL AMONG LISTED AUSTRIAN COMPANIES

3.1. Sample selection and methodology of the research

The literature shows that adjustments of regulatory requirements, innovations in technical reporting standards and new presentation forms of financial reports are predominantly relevant for companies listed at stock markets [21]. For listed companies, financial communications is a core strategic issue, and thus developments in this field are of high relevance. Therefore, the questionnaire survey addresses primarily this group and was designed to demonstrate its perspectives.

The quantitative online survey was conducted among Austrian listed companies from January to February 2016. At the time, the Austrian stock exchange listed a total of 57 companies from which 39 (68%) were listed in ATX Prime, 9 (16%) in the Mid Market and 9 (16%) in the Standard Market. We received a total of 37 responses from which 25 responses were evaluable. Accordingly, the overall response rate was 44%. Since the survey focus results in a relatively small sample size, the methodological approach remains descriptive. The results presented and discussed here should be interpreted in the light of this fact. However, the research findings provide an overview comparable with international research and a basis for further studies.

The questions were derived from extant literature and reflect (1) the current role of financial reporting, the estimated trends in financial reporting, the relevance of technical reporting standards in the companies, and the challenges associated with the new requirements, (2) the diffusion and adoption of XBRL among Austrian listed companies, and (3) the reasons for and against the implementation of XBRL in companies and the promoting and inhibiting factors in this context.

The first section of the questionnaire covered demographic information such as the company size, stock market, industry, working area and management level of the respondents, and the role of financial reporting in the company. Table 1 provides some basic information regarding the sample structure and the frequency distribution in terms of demographic data. The second part of the questionnaire contained seven general questions that cover the expected development of financial reporting in the future (Table 2). Additionally, the respondents were asked how they estimate the relevance of technical reporting standards (Table 3) and which challenges they expect in the context of the implementation of new technical reporting standards (Table 4). The third section of the questionnaire examines the knowledge and adoption of XBRL among the Austrian companies and the level of expertise among the respondents (Table 5). This part is followed by detailed questions that address respondents who know and are more or less familiar with XBRL. This section covered two general questions on reasons for and against the implementation of XBRL and two further questions on advantages and disadvantages associated with the adoption of XBRL.

The following chapter explores the key research findings of the survey, detailing the estimated trends in financial reporting and technical reporting standards and the diffusion and adoption of XBRL among Austrian listed companies.

Table 1: Sample structure

Outline criteria	n (%)
1. Market	
1a. ATX Prime	19 (76.0)
1b. Mid or Standard Market	6 (24.0)
2. Industry	
2a. Basic Industries	6 (24.0)
2b. Industrial Goods & Services	8 (32.0)
2c. Consumer Products	3 (12.0)
2d. Consumer Services	1 (4.0)
2e. Financials	4 (16.0)
2f. Technology & Telecom	2 (8.0)
2g. Utilities	1 (4.0)
3. Working area of respondents	
3a. Investor Relations	19 (76.0)
3b. Public Relations	1 (4.0)
3c. Controlling	3 (12.0)
3d. Misc.	2 (8.0)
4. Management level of respondents	
4a. Top Management	7 (28.0)
4b. Middle Management	9 (36.0)
4c. Lower Management	4 (16.0)
4d. Staff sections	4 (16.0)
4e. Misc.	1 (4.0)
5. Role of FR within the company	
5a. FR is used to fulfil the legal requirements only	3 (12.00)
5b. We plan to make FR an integral component of our communication strategy	5 (20.00)
5c. We established FR as a central component of our communication strategy	17 (68.00)

Note: This table displays the frequencies regarding (1) the market, in which the companies are listed, (2) the industry, in which the companies are active, (3) the working area, (4) the management level of the respondents, and (5) the role of financial reporting within the company.

3.2. Research Findings

3.2.1. Estimated trends in financial reporting and the relevance of technical reporting standards

The first section of the survey investigated the current role of financial reporting in Austrian listed companies. Table 1(5) demonstrates that for 68% of all companies, Financial Reporting (FR) plays a crucial role in the corporate communication and goes far beyond the fulfilment of legal requirements. Further 20% are aware of the strategic relevance of financial reporting and plan to make financial reporting an integral component of the company's communication strategy. Only 12% of the respondents use financial reporting for fulfilling legal requirements only. Thus, for the majority of Austrian listed companies financial reporting is important not only in the communication to investors and regulators, but also to other stakeholders affected by the financial prosperity of a company such as employees, suppliers etc. There is a high level of awareness that financial reporting is a decisive factor in the relations between the company and its environment.

Table 2: Estimated trends in financial reporting

Trend	n Min	Mean Max	Mdn Skewness	SD Kurtosis
Greater need for FR	25	1.96	2.00	0.735
	1	3	0.064	-1.035
Increase frequency	25	2.84	3.00	0.898
	1	4	-0.413	-0.389
New forms of presentation	25	2.20	2.00	0.816
	1	4	0.599	0.362
New forms of narration	25	2.08	2.00	0.759
	1	4	0.483	0.444
Increase personalization	25	2.00	2.00	0.957
	1	4	0.619	-0.485
Increase automation	25	1.88	2.00	0.666
	1	3	0.134	-0.557
Increase standardization	25	1.72	2.00	0.614
	1	3	0.224	-0.445

Note: This table summarises views of all respondents regarding the estimated trends in financial reporting. Means reflects a Likert scale where 1 = fully agree, 2 = somewhat agree, 3 = rather disagree, 4 = disagree. As shown by the skewness and kurtosis, the data is not normally distributed and mirror clear tendencies.

The second section examined the estimated trends in the context of financial reporting in the future. A vast majority of the respondents agree or fully agree that technical standardisation (92%) and automatization (84%) in financial reporting frequency will increase in the future. The need for a higher technical standardisation and automatization could result from the assumption that the need for financial information will increase in general (76%) and will have to be more personalized and target-group oriented (72%) which requires new forms of narration (76%) and presentation (72%) in financial reporting. Thus, managing the higher amount and complexity of financial reporting will be a new challenge for controlling, investor relations, public relations and IT departments. Automatization on top of new technical standards such as XBRL, seem to be the necessary applications to manage these upcoming affordances. Implementation of new technological reporting standards can be entailed with multiple challenges. Table 2 illustrates the corresponding frequency distribution.

Table 3: Challenges of implementation of technical reporting standards

Question	Yes n (%)	No n (%)
What challenges do companies have to face by implementing technical reporting standards? (n=25)		
a. Adjustment of existing workflow and conventions	22 (88.0)	3 (12.0)
b. Education and training of staff in charge	21 (84.0)	4 (16.0)
c. Development of a new policy for the use of financial data	10 (40.0)	15 (60.0)
d. Missing IT expertise	5 (20.0)	20 (80.0)
e. Inestimable follow-up costs	8 (32.0)	17 (68.0)
f. Guarantee of data security	21 (84.0)	4 (16.0)

Note: This table reports the descriptive statistics (frequencies) of challenges associated with the implementation of new technical reporting standards by all respondents independent of their XBRL knowledge (n = 25).

The two main hurdles to the adoption of technical reporting standards seem to be related to staff and processes (Table 3). 88% of all respondents think that the education and training for staff in charge and the need of adjustment of existing workflows and reporting conventions are the two most important challenges. Thus, XBRL might be rather a challenge for HR, organisation and change management than for IT management. Another challenge for a sizeable portion of respondents (80%) is a technical issue concerning the data safety (low data volatility) and data security (controlled accessibility). Inestimable follow-up costs and the development of a new financial data policy seem to concern 36% of all respondents. Missing IT expertise consider 20% of all respondents a challenge.

3.2.2. Diffusion and adoption of XBRL among Austrian listed companies

The third section was dedicated to the diffusion and adoption of XBRL. Generally speaking, 88% of all respondents estimate the relevance of technical reporting standards as high or very high. Table 4 illustrates the corresponding frequency distribution.

Despite the general awareness about the importance of technical reporting standards, the results indicate a poor knowledge of XBRL among Austrian listed firms. As illustrated in Table 5(1), a sizeable portion of the respondents (72%) don't know XBRL at all. Only 7 out of 25 respondents (28%) know XBRL, whereas none of the respondents consider him- or herself an expert. The level of expertise among those who know XBRL is predominantly low (71.4%) or non-existent (14.3%). Only 14.3% describe their level of expertise as middle (Table 5(2)).

Table 4: Estimated relevance of technical reporting standards

	No.	Mean	Mdn	SD
	Min	Max	Skewness	Kurtosis
How do you estimate the relevance of technical reporting standards in the future?				
	24	1.92	2.00	0.504
	1	3	-0.196	1.463

Note: This table reports the views of all respondents who rated the relevance of technical reporting standards in the future from 1 to 4 regarding. Means reflects a Likert scale where 1 = very high, 2 = high, 3 = low, 4 = negligible, 5 = don't know. n = 24 instead of 25, because only scale 1-4 was taken into account.

Considering the adoption of XBRL, the survey shows that XBRL has not been an issue of financial reporting practice at the beginning of 2016 (Table 5(3)). Only one company already reacted to the upcoming challenges and applies the new technical standard (14.3%). 28.6% of the companies are aware of the upcoming challenges and plan to adopt XBRL within the next 5 years. The vast majority of 59.2% is hardly aware of the requirements and possible solutions.

Table 5(3) illustrates that the respondents have neither concrete plans to adopt XBRL for the time being (42.8%) nor state that they will adopt XBRL at all (14.3%). Just one respondent has already adopted XBRL (14.3%). And just two respondents plan to adopt XBRL within the next five years (28.6%).

Table 5: Diffusion and adoption of XBRL

Question	n (%)
1. Do you know XBRL? (n=25)	
1a. Yes	7 (28.00)
1b. No	18 (72.00)
2. What is your level of XBRL expertise? (n=7)	
2a. High	0 (0.00)
2b. Middle	1 (14.30)
2c. Low	5 (71.40)
2d. Non-existent	1 (14.30)
3. To what extent has XBRL been installed in your company? (n=7)	
3a. We already use XBRL	1 (14.30)
3b. We plan to adopt XBRL within the next 5 years	2 (28.60)
3c. We have no plans to adopt XBRL for the time being	3 (42.80)
3d. We won't adopt XBRL	1 (14.30)

Note: This table reports the frequencies regarding (1) the spread of knowledge of XBRL among the respondents, (2) the self estimated level of XBRL expertise among the respondents who know XBRL, and (3) the level of XBRL adoption within the investigated companies knowing XBRL.

Table 6: Reasons for and against the implementation of XBRL

Question	n (%)
1. What were the reasons for the implementation of XBRL? (n=3)	
1a. We deliberately decided to adopt XBRL	0 (0.00)
1b. We were forced to adopt XBRL	0 (0.00)
1c. XBRL came in the course of a technical upgrade	1 (25.00)
1d. XBRL was part new reporting routines	1 (25.00)
1e. Misc. reasons for XBRL adoption	2 (50.00)
2. What were the reasons against the implementation of XBRL? (n=4)	
2a. No need for XBRL	0 (0.00)
2b. We use other standards (e.g. Edifact, ebXML)	0 (0.00)
2c. XBRL is no issue	4 (66.67)
2d. Implementation costs	1 (16.67)
2e. Immaturity of the technology	1 (16.67)
2f. Missing expertise	0 (0.00)
2g. Security issues	0 (0.00)
2h. Misc. reasons against XBRL adoption	0 (0.00)

Note: This table displays the frequencies regarding (1) the reasons for and (2) reasons against the implementation of XBRL among respondents who (1) know XBRL and has already adopted or plan to adopt XBRL within the next 5 years (n = 3) and (2) know XBRL and have no plans to adopt XBRL (n = 4).

From the latter three respondents, no company adopted XBRL deliberately (Table 6(1)). If XBRL was adopted, then as part of new reporting routines or in the course of technological upgrades. The intention to improve financial reporting to and communication with stakeholders doesn't seem to have played a role at all. Thus, the adoption of XBRL does not seem to be the result of a new communication culture, but rather a technical issue. Despite the low adoption rate, no special inhibiting reasons could be identified (Table 6(2)). XBRL is rather not an issue at all (66.67%) or doesn't seem to be a mature technology (16.67%).

3.2.3. Awareness of the benefits and barriers of the adoption of XBRL

In the last phase we compared the perceived benefits and obstacles of XBRL between those respondents who have adopted or plan to adopt XBRL, and those respondents who know XBRL but have no concrete adoption strategy yet (Table 7).

The three respondents who know XBRL replied that reusability and comparability of financial data, higher flexibility and analytical capabilities, and decrease of processing errors are seen as the main advantages of XBRL. Further benefits of the new technology that were recognized by the respondents are decrease of reporting costs, improved data portability between data systems, improved findability of the data, acceleration of data processing and reporting processes, and miscellaneous. Trustworthiness of the data source or improved data portability between data systems are not considered an advantage at all. A considerable portion of respondents see the additional costs to occur as the main drawback, whereas for some the implementation costs are expected to be the greatest strain, followed by the cost of XBRL-software and additional training costs for employees. Further disadvantages seen by the respondents are security issues, complexity of XBRL and disruption of reporting routines. Missing software tools or volatility of XBRL are not seen as disadvantages at all. 25% of respondents see also other, not specific aspects as disadvantages of XBRL.

Table 7: Benefits and obstacles of XBRL implementation

	Know and adopted XBRL (n=3)		Know, but have not (yet) adopted XBRL (n=4)	
	Yes	No	Yes	No
1. What are the benefits of XBRL?				
1a. Reusability of financial data	2	1	1	3
1b. Comparability of financial data	1	2	2	2
1c. Acceleration of data processing	0	3	0	4
1d. Higher flexibility and analytical capabilities	2	1	1	3
1e. Improved findability of financial data	1	2	1	3
1f. Improved data portability between IT systems	1	2	1	3
1g. Improved cross-system integrity of data	0	3	0	4
1h. Trustworthiness of the data source	0	3	0	4
1i. Decrease of reporting costs	2	1	0	4
1j. Decrease of processing errors	2	1	1	3
1k. Misc.	1	2	1	3
2. What are the obstacles hindering the adoption of XBRL?				
2a. Additional training for employees	1	2	2	2
2b. Implementation costs	2	1	3	1
2c. Disruption of reporting routines	1	2	0	4
2d. Costs for XBRL software	1	2	3	1
2e. Complexity of standards	0	3	1	3
2f. Volatility of standards	0	3	0	4
2g. Missing software tools	0	3	0	4
2h. Security issues	0	3	1	3
2i. Misc.	2	1	2	2

Note: This table shows the (1) advantages and (2) disadvantages of XBL estimated by respondents who know XBRL and already adopted or plan to adopt XBRL within the next 5 years (n = 3) and know XBRL and have no plans to adopt XBRL (n = 4).

On the contrary, the four respondents with no specific adoption plans perceive the comparability of financial data as the main benefit. Obstacles are implementation costs, additional training demand for employees and costs for XBRL software.

4. DISCUSSION AND CONCLUSION

The survey results correspond with the findings of research conducted in the US and other European countries in the recent years. The situation among Austrian listed companies doesn't differ significantly from other countries and stock markets. Moreover, the survey results confirm the general lack of knowledge about XBRL which stands in contradiction to the great awareness for the need of target-group oriented financial reporting and high relevance of technical reporting standards in the future. This finding is surprising and worrying with respect to the length of time XBRL has been a topic of discussion among researchers and governmental and professional entities. Only one third of all respondents know XBRL, whereas XBRL has been a topic of the AICPA, the SEC, the IASB, and other major entities since 2004 and experts think that we reached the tipping point toward the use of XBRL [12]. That leaves the impression that the discourse in the previous years failed to reach the Austrian companies.

Another fact confirmed by the survey is that private initiatives to implement XBRL hardly exist and can't be expected. If new information technologies should be adopted for more accurate, reliable and customized financial reporting, external initiatives seem to be necessary to enhance the adoption of XBRL in private companies.

Neglecting the demand for new reporting standards with respect to customized financial information provided by new technical standards such as XBRL might weaken a company's position in the stock market and in the public perception. The high share of international investors in the Austrian stock market might even amplify the negative aspects on not adopting XBRL and a new reporting culture. International investors compare reporting standards in an international context and tend to prefer companies and stock markets that answer investors' and stakeholders' demand for new financial reporting standards. However, reacting to these new affordances means in the current environment a strategic advantage and could strengthen the company's position and enhance its value.

REFERENCES

[1] Baldwin, A. A., & Trinkle, B. S. (2011). The Impact of XBRL: A Delphi Investigation. The International Journal of Digital Accounting Research, 11. http://doi.org/10.4192/1577-8517-v11_1

[2] Debreceny, R. et al. (2009). XBRL for interactive data: engineering the information value chain. London ; New York: Springer.

[3] Dhole, S. et al. (2015). Effects of the SEC's XBRL mandate on financial reporting comparability. International Journal of Accounting Information

Systems, 19, p. 29–44. http://doi.org/10.1016/j.accinf.2015.11.002

[4] Dunne, T. et al. (2013). Stakeholder engagement in internet financial reporting: The diffusion of XBRL in the UK. The British Accounting Review, 45(3), p. 167–182. http://doi.org/10.1016/j.bar.2013.06.012

[5] ERICA (2010). XBRL in European CBSO. Document nº 6. III WORKING GROUP ON IFRS IMPACT AND CBSO DATABASES. Innsbruck, 7th – 8th October 2010. https://www.nbb.be/doc/ba/xbrl/pub/2010_10_wgiii_xbrl.pdf

[6] Escobar-Rodriguez, T.; Gago-Rodriguez, S. (2010). “We were the first to support a major innovation". Research into the motivations of Spanisch pioneers in XBRL. Revista de Contabilidad, 15(1), p. 91-108. http://www.redalyc.org/articulo.oa?id=359733642003 http://doi.org/10.1016/j.bar.2013.06.012

[7] García, R., & Gil, R. (2010). Linking XBRL Financial Data. In D. Wood, Linking Enterprise Data , p. 103–125. Boston, MA: Springer US. Abgerufen von http://link.springer.com/10.1007/978-1-4419-7665-9_6

[8] Graning, A., Felden, C., & Piechocki, M. (2011). Status Quo and Potential of XBRL for Business and Information Systems Engineering. Wirtschaftsinformatik, 53(4), p. 225–234. http://doi.org/10.1007/s11576-011-0282-2 http://link.springer.com/10.1007/978-1-4419-7665-9_6

[9] Guilloux, V., Locke, J., & Lowe, A. (2013). Digital business reporting standards: mapping the battle in France. European Journal of Information Systems, 22(3), p. 257–277. http://doi.org/10.1057/ejis.2012.5

[10] ICAEW (2011). The future of XBRL in Europe: Impetus, institutions and interrelationships. Workshop on the future of XBRL in Europe. ICAEW and University of Birmingham (ISARG), 25th January 2011. http://www.icaew.com/~/media/corporate/files/about%20icaew/what%20we%20do/thought%20leadership/the%20future%20of%20xbrl%20in%20europe%20final%20summary%20for%20release.ashx

[11] Kernan, K. (2008). XBRL Around the World. In: Journal of Accountancy, October 1, 2008. http://www.journalofaccountancy.com/issues/2008/oct/xbrlaroundtheworld.html http://doi.org/10.1057/ejis.2012.5

[12] Lester, W. F. (2007). XBRL: The New Language of Corporate Financial Reporting. Business Communication Quarterly, 70(2), p. 226–231. http://doi.org/10.1177/10805699070700020603

[13] Liu, C., Wang, T., & Yao, L. J. (2014). XBRL’s impact on analyst forecast behavior: An empirical study. Journal of Accounting and Public Policy, 33(1), p. 69–82. http://doi.org/10.1016/j.jaccpubpol.2013.10.004

[14] Melancon, B. (2006). Testimony of Barry Melancon, President and CEO, AICPA, before the Capital Markets, Insurance and Government Sponsored Enterprises Subcommittee of the House Committee on Financial Services Concerning Fostering Accuracy and Transparency in Financial Reporting. Retrieved February 15, 2016 from http://financialservices.house.gov/media/pdf/032906bm.pdf

[15] Oades, C. (2008). Information management challenges for the professional accountant in business. Business Information Review, 25(3), p. 160–164. http://doi.org/10.1177/0266382108095041

[16] O’Riain, S., Curry, E., & Harth, A. (2012). XBRL and open data for global financial ecosystems: A linked data approach. International Journal of Accounting Information Systems, 13(2), p. 141–162. http://doi.org/10.1016/j.accinf.2012.02.002

[17] Rodriguez Bolivar, M. P., Caba Perez, C., & Lopez Hernandez, A. M. (2007). E-Government and Public Financial Reporting: The Case of Spanish Regional Governments. The American Review of Public Administration, 37(2), p. 142–177. http://doi.org/10.1177/0275074006293193

[18] Rodriguez Bolivar, M. P. (2009). Evaluating Corporate Environmental Reporting on the Internet: The Utility and Resource Industries in Spain. Business & Society, 48(2), p. 179–205. http://doi.org/10.1177/0007650307305370

[19] SEC (2009). Interactive Data to Improve Financial Reporting. Final rule. RIN 3235-AJ71. April 13, 2009. https://www.sec.gov/rules/final/2009/33-9002.pdf

[20] Sinnett, William M. (2013). SEC REPORTING AND THE IMPACT OF XBRL: 2013 SURVEY. Financial Executives Research Foundation (FERF). http://www.financialexecutives.org/ferf/download/2013%20Final/2013-022.pdf

[21] Zitzmann, A., Fischer, T., Decker, T. (2009). Rechtsfragen der IR. In: Kirchhoff, Klaus; Piwinger, Manfred (Hg.). Praxishandbuch Investor Relations. Wiesbaden: Gabler Verlag, p. 93-13

Comparing Shallow versus Deep Neural Network Architectures for Automatic Music Genre Classification

Alexander Schindler
Austrian Institute of Technology
Digital Safety and Security
Vienna, Austria
alexander.schindler@ait.ac.at

Thomas Lidy, Andreas Rauber
Vienna University of Technology
Institute of Software Technology
Vienna, Austria
lidy,rauber@ifs.tuwien.ac.at

Abstract

In this paper we investigate performance differences of different neural network architectures on the task of automatic music genre classification. Comparative evaluations on four well known datasets of different sizes were performed including the application of two audio data augmentation methods. The results show that shallow network architectures are better suited for small datasets than deeper models, which could be relevant for experiments and applications which rely on small datasets. A noticeable advantage was observed through the application of data augmentation using deep models. A final comparison with previous evaluations on the same datasets shows that the presented neural network based approaches already outperform state-of-the-art handcrafted music features.

1 Introduction

Music classification is a well researched topic in Music Information Retrieval (MIR) [FLTZ11]. Generally, its aim is to assign one or multiple labels to a sequence or an entire audio file, which is commonly accomplished in two major steps. First, semantically meaningful audio content descriptors are extracted from the sampled audio signal. Second, a machine learning algorithm is applied, which attempts to discriminate between the classes by finding separating boundaries in the multidimensional feature-spaces. Especially the first step requires extensive knowledge and skills in various specific research areas such as audio signal processing, acoustics and/or music theory. Recently many approaches to MIR problems have been inspired by the remarkable success of Deep Neural Networks (DNN) in the domains of computer vision [KSH12], where deep learning based approaches have already become the *de facto standard*. The major advantage of DNNs are their *feature learning* capability, which alleviates the domain knowledge and time intensive task of crafting audio features by hand. Predictions are also made directly on the modeled input representations, which is commonly raw input data such as images, text or audio spectrograms. Recent accomplishments in applying Convolutional Neural Networks (CNN) to audio classification tasks have shown promising results by outperforming conventional approaches in different evaluation campaigns such as the Detection and Classification of Acoustic Scenes and Events (DCASE) [LS16a] and the Music Information Retrieval Evaluation EXchange (MIREX) [LS16b].

An often mentioned paradigm concerning neural networks is that deeper networks are better in modeling non-linear relationships of given tasks [SLJ+15]. So far preceding MIR experiments and approaches reported in literature have not explicitly demonstrated the advantage of deep over shallow network architectures in a magnitude similar to results reported from the computer vision domain. This may be related to the absence of similarly large datasets as they are available in the visual related research areas. A special focus of this paper is thus set on the performance of neural networks on small datasets, since data availability is still a problem in MIR, but also because many tasks involve the processing of small collections. In this paper we present a performance evaluation of shallow and deep neural network architectures. These models and the applied method will be detailed in Section 2. The evaluation will be performed on well known music genre classification datasets in the domain of Music Information Retrieval. These datasets and the evaluation procedure will be described in Section 3. Finally we draw conclusions from the results in Section 5 and give an outlook to future work.

In: W. Aigner, G. Schmiedl, K. Blumenstein, M. Zeppelzauer (eds.): Proceedings of the 9th Forum Media Technology 2016, St. Pölten, Austria, 24-11-2016, published at http://ceur-ws.org

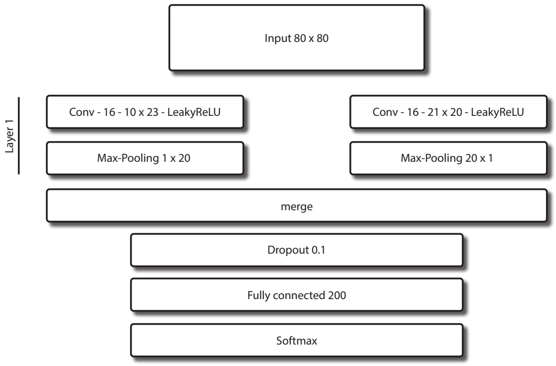


Figure 1: Shallow CNN architecture

2 Method

The parallel architectures of the neural networks used in the evaluation are based on the idea of using a time and a frequency pipeline described in [PLS16], which was successfully applied in two evaluation campaigns [LS16a, LS16b]. The system is based on a parallel CNN architecture where separate CNN Layers are optimized for processing and recognizing music relations in the frequency domain and to capture temporal relations (see Figure 1).

The Shallow Architecture: In our adaption of the CNN architecture described in [PLS16] we use two similar pipelines of CNN Layers with 16 filter kernels each followed by a Max Pooling layer (see Figure 1). The left pipeline aims at capturing frequency relations using filter kernel sizes of 10×23 and Max Pooling sizes of 1×20. The resulting 16 vertical rectangular shaped feature map responses of shape 80×4 are intended to capture spectral characteristics of a segment and to reduce the temporal complexity to 4 discrete intervals. The right pipeline uses a filter of size 21×20 and Max Pooling sizes of 20×1. This results in horizontal rectangular shaped feature maps of shape 4×80. This captures temporal changes in intensity levels of four discrete spectral intervals. The 16 feature maps of each pipeline are flattened to a shape of 1×5120 and merged by concatenation into the shape of 1×10240, which serves as input to a 200 units fully connected layer with a dropout of 10%.

The Deep Architecture: This architecture follows the same principles of the shallow approach. It uses a parallel arrangement of rectangular shaped filters and Max-Pooling windows to capture frequency and temporal relationships at once. But, instead of using the information of the large feature map responses, this architecture applies additional CNN and pooling layer pairs (see Figure 2). Thus, more units can be applied to train on the subsequent smaller input feature maps. The first level of the parallel layers are similar to the original approach. They use filter kernel sizes of 10×23 and 21×10 to capture frequency and temporal relationships. To retain these characteristics the sizes of the convolutional filter kernels as well as the feature maps are sub-sequentially divided in halves by the second and third layers. The filter and Max Pooling sizes of the fourth layer are slightly adapted to have the same rectangular shapes with one part being rotated by 90°. As in the shallow architecture the same sizes of the final feature maps of the parallel model paths balances their influences on the following fully connected layer with 200 units with a 25% dropout rate.

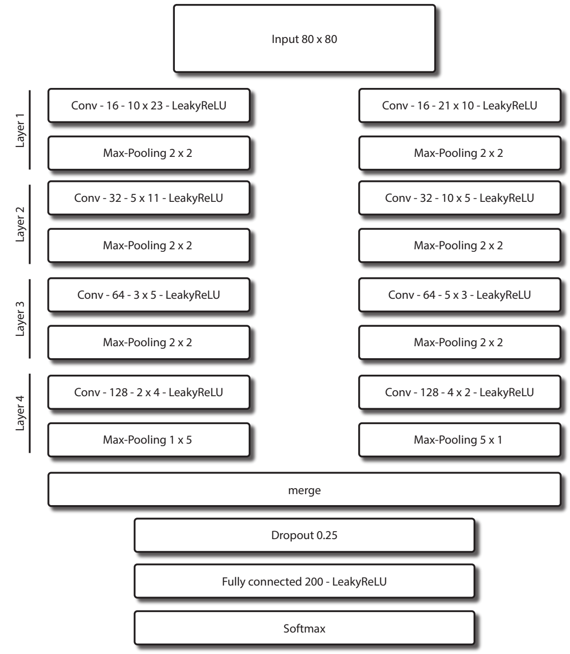


Figure 2: Deep CNN architecture

2.1 Training and Predicting Results

In each epoch during training the network multiple training examples sampled from the segment-wise log-transformed Mel-spectrogram analysis of all files in the training set are presented to both pipelines of the neural network. Each of the parallel pipelines of the architectures uses the same 80 × 80 log-transformed Mel-spectrogram segments as input. These segments have been calculated from a fast Fourier transformed spectrogram using a window size of 1024 samples and an overlap of 50% from 0.93 seconds of audio transformed subsequently into Mel scale and Log scale.. For each song of a dataset 15 segments have been randomly chosen.

All trainable layers used the *Leaky ReLU* activation function [MHN13], which is an extension to the ReLU (Rectifier Linear Unit) that does not completely cut off activation for negative values, but allows for negative values close to zero to pass through. It is defined by adding a coefficient α in $f(x) = \alpha x$, for $x < 0$, while keeping $f(x) = x$, for $x \geq 0$ as for the ReLU. In our architectures, we apply Leaky ReLU activation with $\alpha = 0.3$. L_1 weight regularization with a penalty of 0.0001 was applied to all trainable parameters. All networks were trained towards *categorical-crossentropy* objective using the stochastic *Adam* optimization [KB14] with $beta_1 = 0.9$, $beta_2 = 0.999$, $epsilon = 1e{-}08$ and a learning rate of 0.00005.

The system is implemented in Python and using *librosa*

[MRL+15] for audio processing and Mel-log-transforms and *Theano*-based library *Keras* for Deep Learning.

2.1.1 Data Augmentation

To increase the number of training instances we experiment with two different audio data augmentation methods. The deformations were applied directly to the audio signal preceding any further feature calculation procedure described in Section 2.1. The following two methods were applied using the MUDA framework for musical data augmentation [MHB15]:

Time Stretching: slowing down or speeding up the original audio sample while keeping the same pitch information. Time stretching was applied using the multiplication factors 0.5, 0.2 for slowing down and 1.2, 1.5 for increasing the tempo.

Pitch Shifting: raising or lowering the pitch of an audio sample while keeping the tempo unchanged. The applied pitch shifting lowered and raised the pitch by 2 and 5 semitones.

For each deformation three segments have been randomly chosen from the audio content. The combinations of the two deformations with four different factors each resulted thus in 48 additional data instances per audio file.

3 Evaluation

As our system analyzes and predicts multiple audio segments per input file, there are several ways to perform the final prediction of an input instance:

Raw Probability: The raw accuracy of predicting the segments as separated instances ignoring their file dependencies.

Maximum Probability: The output probabilities of the Softmax layer for the corresponding number of classes of the datasets are summed up for all segments belonging to the same input file. The predicted class is determined by the maximum probability among the classes from the summed probabilities.

Majority Vote: Here, the predictions are made for each segment processed from the audio file as input instance to the network. The class of an audio segment is determined by the maximum probability as output by the Softmax layer for this segment instance. Then, a majority vote is taken on all predicted classes from all segments of the same input file. Majority vote determines the class that occurs most often.

We used stratified 4-fold cross validation. Multi-level stratification was applied paying special attention to the multiple segments used per file. It was ensured that the files were distributed according their genre distributions and that no segments of a training file was provided in the corresponding test split.

The experiments were grouped according to the four different datasets. For each dataset the performances for the shallow and deep architecture were evaluated followed by the experiments including data augmentation. The architectures were further evaluated according their performance after a different number of training epochs. The networks were trained and evaluated after 100 and 200 epochs without early stopping. Preceding experiments showed that test accuracy could improve despite rising validation loss though on smaller sets no significant improvement was recognizable after 200 epochs. For the experiments with data augmentation, the augmented data was only used to train the networks (see Table 3.1). For testing the network the original segments without deformations were used.

Data	Tracks	cls	Train wo. au.	Train w. au.	Test
GTZAN	1,000	10	11,250	47,250	3,750
ISMIR G.	1,458	6	16,380	68,796	5,490
Latin	3,227	10	36,240	152,208	12,165
MSD	49,900	15	564,165	—	185,685

Table 1: Overview of the evaluation datasets, their number of classes (cls) and their corresponding number of test and training data instances without (wo. au.) and with (w. au.) data augmentation.

3.1 Data Sets

For the evaluation four data sets have been used. We have chosen these datasets due to their increasing number of tracks and because they are well known and extensively evaluated in the automatic genre classification task. This should also provide comparability with experiments reported in literature.

GTZAN: This data set was compiled by George Tzanetakis [Tza02] in 2000-2001 and consists of 1000 audio tracks equally distributed over the 10 music genres: blues, classical, country, disco, hiphop, pop, jazz, metal, reggae, and rock.

ISMIR Genre: This data set has been assembled for training and development in the ISMIR 2004 Genre Classification contest [CGG+06]. It contains 1458 full length audio recordings from Magnatune.com distributed across the 6 genre classes: Classical, Electronic, JazzBlues, MetalPunk, RockPop, World.

Latin Music Database (LMD): [SKK08] contains 3227 songs, categorized into the 10 Latin music genres Axé, Bachata, Bolero, Forró, Gaúcha, Merengue, Pagode, Salsa, Sertaneja and Tango.

Million Song Dataset (MSD): [BMEWL11] a collection of one million music pieces, enables methods for large-scale applications. It comes as a collection of metadata such as the song names, artists and albums, together with a set of features extracted with the The Echo Nest services, such as loudness, tempo, and MFCC-like features. We used the CD2C genre assignments as ground truth [Sch15] which are an adaptation of the MSD genre label assignments presented in [SMR12]. For the experiments a sub-set of approximately 50.000 tracks was sub-sampled.

4 Results

The results of the experiments are provided in Table 4. For each dataset all combinations of experimental results were tested for significant difference using a Wilcoxon signed-rank test. None of the presented results showed a significant difference for $p < 0.05$. Thus, we tested at the next higher level $p < 0.1$. The following observations on the datasets were made:

GTZAN: Training the models with 200 epochs instead of only 100 epochs significantly improved the *raw* and *max* accuracies for the shallow models. An additional test on training 500 epochs showed no further increase in accuracy for any of the three prediction methods. Training longer had no effect on the deep model due to early over-fitting. No significant differences were observed between shallow and deep models except for the raw prediction values of the shallow model (200 epochs) exceeding those of the deep model (200 epochs). While the improvements through data augmentation on deep models compared to the un-augmented longer trained deep models are not significant, considerable improvements of 4.2% were observed for models trained for the same number of epochs. An interesting observation is the negative effect of data augmentation on the shallow models where longer training outperformed augmentation.

ISMIR Genre: Training more epochs only had a significant positive effect on the *max* and *maj* values of the deep model but none for the shallow ones. The deep models showed no significant advantage over the shallow architectures which also showed higher *raw* prediction values even on shorter trained models. Data augmentation improved the predictions of both architectures with significant improvements for the *raw* values. Especially the deep models significantly profited from data augmentation with *max* values increased by 3.08% for models trained for the same number of epochs and 2.05% for the longer trained models. The improvements from deep over shallow models using augmented data were only significant for the *raw* values.

Latin: Training more epochs only had a positive effect for the *raw* and *max* values of the shallow model, but not for the deep architecture. On this dataset, the deep model significantly outperformed the shallow architecture including the shallow model trained using data augmentation. Data augmentation improved the significantly improved the performance of the deep models by 1,61% for the *max* values. Similar to the GTZAN dataset, data augmentation showed a degrading effect on the shallow model which showed significantly higher accuracy values by training for more epochs.

MSD: A not significant advantage of deep over shallow models was observed. Experiments using data augmentation and longer training were omitted due to the already large variance provided by the MSD which multiplies the preceding datasets by factors from 15 to 50.

D	Model	raw	max	maj	ep
GTZAN	shallow	66.56 (0.69)	78.10 (0.89)	77.80 (0.52)	100
	deep	65.69 (1.23)	78.60 (1.97)	78.00 (2.87)	100
	shallow	67.49 (0.39)	80.80 (1.67)	80.20 (1.68)	200
	deep	66.19 (0.55)	80.60 (2.93)	80.30 (2.87)	200
	shallow aug	66.77 (0.78)	78.90 (2.64)	77.10 (1.19)	100
	deep aug	**68.31** (2.68)	**81.80** (2.95)	**82.20** (2.30)	100
ISMIR Genre	shallow	75.66 (1.30)	85.46 (1.87)	84.77 (1.43)	100
	deep	74.53 (0.52)	84.08 (1.13)	83.95 (0.97)	100
	shallow	75.43 (0.65)	84.91 (1.96)	85.18 (1.27)	200
	deep	74.51 (1.71)	85.12 (0.76)	85.18 (1.23)	200
	shallow aug	76.61 (1.04)	86.90 (0.23)	86.00 (0.54)	100
	deep aug	**77.20** (1.14)	**87.17** (1.17)	**86.75** (1.41)	100
Latin	shallow	79.80 (0.95)	92.44 (0.86)	92.10 (0.97)	100
	deep	81.13 (0.64)	94.42 (1.04)	94.30 (0.81)	100
	shallow	80.64 (0.83)	93.46 (1.13)	92.68 (0.88)	200
	deep	81.06 (0.51)	95.14 (0.40)	94.83 (0.53)	200
	shallow aug	78.09 (0.68)	92.78 (0.88)	92.03 (0.81)	100
	deep aug	**83.22** (0.83)	**96.03** (0.56)	**95.60** (0.58)	100
MSD	shallow	58.20 (0.49)	63.89 (0.81)	63.11 (0.74)	100
	deep	**60.60** (0.28)	**67.16** (0.64)	**66.41** (0.52)	100

Table 2: Experimental results for the evaluation datasets (D) at different number of training epochs (ep): Mean accuracies and standard deviations of the 4-fold cross-evaluation runs calculated using raw prediction scores (raw) and the file based maximum probability (max) and majority vote approach (maj).

5 Conclusions and Future Work

In this paper we evaluated shallow and deep CNN architectures towards their performance on different dataset sizes in music genre classification tasks. Our observations showed that for smaller datasets shallow models seem to be more appropriate since deeper models showed no significant improvement. Deeper models performed slightly better in the presence of larger datasets, but a clear conclusion that deeper models are generally better could not be drawn. Data augmentation using time stretching and pitch shifting significantly improved the performance of deep models. For shallow models on the contrary it showed a negative effect on the small datasets. Thus, deeper models should be considered when applying data augmentation. Comparing the presented results with previously reported evaluations on the same datasets [SR12] shows, that the CNN based approaches already outperform handcrafted music features such as the Rhythm Patterns (RP) family [LSC+10] (highest values: GTZAN 73.2%, ISMIR Genre 80.9%, Latin 87.3%) or the in the referred study presented Temporal Echonest Features [SR12] (highest values: GTZAN 66.9%, ISMIR Genre 81.3%, Latin 89.0%).

Future work will focus on further data augmentation methods to improve the performance of neural networks on small datasets and the Million Song Dataset as well as on different network architectures.

References

[BMEWL11] Thierry Bertin-Mahieux, Daniel PW Ellis, Brian Whitman, and Paul Lamere. The million song dataset. In *ISMIR*, volume 2, page 10, 2011.

[CGG+06] Pedro Cano, Emilia Gómez, Fabien Gouyon, Perfecto Herrera, Markus Koppenberger, Beesuan Ong, Xavier Serra, Sebastian Streich, and Nicolas Wack. ISMIR 2004 audio description contest. Technical report, 2006.

[FLTZ11] Zhouyu Fu, Guojun Lu, Kai Ming Ting, and Dengsheng Zhang. A survey of audio-based music classification and annotation. *Multimedia, IEEE Transactions on*, 13(2):303–319, 2011.

[KB14] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2014.

[KSH12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[LS16a] Thomas Lidy and Alexander Schindler. CQT-based convolutional neural networks for audio scene classification. In *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2016 Workshop (DCASE2016)*, pages 60–64, September 2016.

[LS16b] Thomas Lidy and Alexander Schindler. Parallel convolutional neural networks for music genre and mood classification. Technical report, Music Information Retrieval Evaluation eXchange (MIREX 2016), August 2016.

[LSC+10] Thomas Lidy, Carlos N. Silla, Olmo Cornelis, Fabien Gouyon, Andreas Rauber, Celso A. A. Kaestner, and Alessandro L. Koerich. On the suitability of state-of-the-art music information retrieval methods for analyzing, categorizing, structuring and accessing non-western and ethnic music collections. *Signal Processing*, 90(4):1032–1048, 2010.

[MHB15] Brian McFee, Eric J Humphrey, and Juan P Bello. A software framework for musical data augmentation. In *International Society for Music Information Retrieval Conference (ISMIR)*, 2015.

[MHN13] Andrew L. Maas, Awni Y. Hannun, and Andrew Y. Ng. Rectifier nonlinearities improve neural network acoustic models. *ICML 2013*, 28, 2013.

[MRL+15] Brian McFee, Colin Raffel, Dawen Liang, Daniel PW Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto. librosa: Audio and music signal analysis in python. In *Proceedings of the 14th Python in Science Conference*, 2015.

[PLS16] Jordi Pons, Thomas Lidy, and Xavier Serra. Experimenting with musically motivated convolutional neural networks. In *Proceedings of the 14th International Workshop on Content-based Multimedia Indexing (CBMI 2016)*, Bucharest, Romania, June 2016.

[Sch15] Hendrik Schreiber. Improving genre annotations for the million song dataset. In *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR 2012)*, Malaga, Spain, 2015.

[SKK08] C N Silla Jr., Celso A A Kaestner, and Alessandro L Koerich. The Latin Music Database. In *Proceedings of the 9th International Conference on Music Information Retrieval*, pages 451—-456, 2008.

[SLJ+15] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–9, 2015.

[SMR12] Alexander Schindler, Rudolf Mayer, and Andreas Rauber. Facilitating comprehensive benchmarking experiments on the million song dataset. In *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR 2012)*, pages 469–474, Porto, Portugal, October 8-12 2012.

[SR12] Alexander Schindler and Andreas Rauber. Capturing the temporal domain in echonest features for improved classification effectiveness. In *Adaptive Multimedia Retrieval*, Lecture Notes in Computer Science, Copenhagen, Denmark, October 24-25 2012. Springer.

[Tza02] G. Tzanetakis. *Manipulation, analysis and retrieval systems for audio signals.* PhD thesis, 2002.

Session 2: Usability and Mobile Applications

Location-Based Learning Games Made Easy

Simon Reinsperger,[1] Florian Grassinger,[1] Iosif Miclaus,[1]
Grischa Schmiedl,[1] Birgit Schmiedl[2], Kerstin Blumenstein[1]

[1]St. Poelten University of Applied Sciences, Austria
[2]HLW Pressbaum, Austria

Abstract

The state of technology is evolving each year, computers get faster and smaller and more and more parts of society get enhanced with it. Education is one of these areas that could benefit from the progress made in technology. This paper explores the possibilities and benefits of mobile location-based learning games and tools for creating such applications. We built a web platform for creating location-based content. The development process and design considerations to expand the functionality for creating context-based games will be documented in this paper. Finally, the results of two user tests will be presented. Both tests were conducted with primary school children to focus on use cases in the educational system.

1 Introduction

The use of technology for educational purposes is not a new idea. However, it definitely has more to offer than what it is used for right now. Almost every student uses smartphones and at the age of 4 already 75% of all children carry such a high-end computer in their pockets [KIND+15]. The possibilities to use these devices in school do not end with playing casual games during a break. The field of mobile computing and especially mobile learning is more and more investigated and context-aware computing is a big opportunity to develop and interact with software. Particularly, location-based applications promise an appealing approach for creating interesting learning experiences. For example, the success of the mobile game "Pokemon Go"[1] shows the fascination location-aware applications have on people, regardless of gender, age or education [Tak16]. The same enthusiasm could be used to enhance learning experiences.

This paper investigates how such a tool could look like and which requirements, use cases and limits of such applications exist. The method consists of a literature research to investigate the current state of location-based learning games (Section 2) and tools for creating such games (Section 3). The second part consists of the development of a prototype as a proof of concept and the evaluation through user tests (Section 4 and 5). Finally, we conclude our work in Section 6 and outline future work.

2 Location-Based Learning Games

Location-based games require players to move through the physical world in order to achieve certain game objectives. Real objects or locations get connected with virtual information, which is accessible through the players mobile device. This kind of games is very flexible concerning the content and is predestined for educational use cases [SYOA+14]. Location-based systems do not necessarily require location-awareness. There are multiple examples of location-based learning applications which do not monitor or react to the users' physical location due to the available hardware at that time. Nowadays, most projects work location-aware, in the sense that the system knows the current location of its user [BBS+10]. All three following examples use GPS (global positioning system) for locating the users. There are other possibilities as well, e.g., tracking the position via WiFi signal strength [CT09] or using QR code and RFID to confirm a location [TKK+09].

In: W. Aigner, G. Schmiedl, K. Blumenstein, M. Zeppelzauer (eds.): Proceedings of the 9th Forum Media Technology 2016, St. Pölten, Austria, 24-11-2016, published at http://ceur-ws.org

[1]`http://www.pokemon.com/us/pokemon-video-games/pokemon-go/`, accessed 26.07.2016.

2.1 Business Education

Puja & Parsons [PP11] explored the possibilities for using a location-based mobile game for teaching college students about business consulting. The game is played in teams. The players goal is to analyze a virtual company, as if they were business consultants and form a recommendation for the company based on their findings and conclusions. The students have to find different virtual interview partners each representing another infrastructural part of the company. These interview partners are bound to certain locations on the campus. Once the players are in a certain range to these locations the interview text appears on the device. After all locations are found and the students have obtained all accessible information, they prepare a final presentation about their findings and analysis. The user test showed some issues with the user interface e.g., difficulties in finding their orientation with only a dot on an abstracted map. Another discovery was that they actually spent little time reading the interviews and documents. The players rated the progress bar as very useful, as a reference of their success. However, this resulted in consuming more attention from the participants than anticipated.

2.2 History Learning

Wake & Baggetun [WB09] developed a location-based mobile learning game to teach students about local history in Bergen, Norway. The players take the role of Premier Lieutenant Bielke, who managed the defense of Bergen during the Napoleonic wars (1803-1815), and visit historic sites in the town to learn more about this historical episode. This concept follows the idea of place-based education, where problems and learning opportunities are based on the students' own surroundings from that applying the learned information on a global perspective. The game is played in teams, which are competing for the lowest score (finishing the game as fast with as few hints provided by the system as possible). A user test with three teams of three people was conducted, each team sharing one GPS-capable mobile device. The participants reported a clear understanding of the game and knew how to use the system. During the questionnaire, the authors found that the distance meter was a crucial tool for the success of the game, the map and hints were rated lower.

2.3 Understanding of Animal Behavior

Savannah [FJS+04] is a game where children (aged 11 to 12) can take the role of a lion in the wild. The game is a rather complex role-playing simulation, where the players can move freely in a predefined area ($100m \times 50m$) and decide their own actions. Basically, the game is about surviving as a lion, which includes hunting, drinking, staying away from trouble, managing their energy, etc. Each player carries a PDA with GPS functionality and headphones. The devices transform the playing field into the savanna by showing images and playing sounds according to the players' position. The device also receives and displays messages like "you are too hot", "you are hungry" or "you are dead - return to the Den". These messages let the players know about their state in the game.

2.4 Insights from the Examples

Looking at these examples, some insights can be obtained about the current state of location-based learning games. The use cases are very broad, the examples reach from games for children to college education and the fields of interests were historical, biological and economical. Thus, location-based learning games are unbiased regarding target audience and content.

The structure of these games often share certain similarities, like the process of going to a location, retrieving information about the next location once the players' presence is confirmed and so on. Overall, all user tests confirm a positive effect on the motivation of the players and their learning experience.

3 Requirements

Many location-based learning applications follow the same structure. The user has to go to a certain place to acquire some new information, based on that they have to find the next location. Having such a repetitive pattern makes it easy to create tools to manage the content of such a learning experience or even create new ones without having programming skills. This kind of application could, for example, be used by teachers to create interactive, location-based content for their classes, benefiting from all the previous mentioned advantages, without the need for expensive projects to create individual location-based learning games for only one use case [SYOA+14]. Keeping such content up-to-date is another advantage of having an underlying - easy to use - system [WHC+06]. Of course, the individuality of game features developed with such a tool will not be guaranteed and complex applications like "Savanna" [FJS+04] can not be realized with such a generic structure.

Weal et al. [WHC+06] worked with teachers and curators and defined requirements for authoring tools of location-based content as follows:

- The process has to fit existing practices, must be fast, simple and achievable in-between daily chores and when new ideas come to mind.

- In-situ authoring should be provided, esp., if the mobile experience created is essentially connected to the real environment. The content providers will benefit from the possibility of creating the content exactly where it will be consumed.
- It should always be possible to go back and refine the already created content. This kind of work does not have to be in-situ because it will probably be more time consuming and reflective than a spontaneous note. This might even profit from having to use another working environment like a desktop computer to get another view.

3.1 In-Situ Authoring

In-situ authoring allows the creation of content directly in the situation when being on sight on a mobile device. The other approach would be to require some kind of static desktop application. Weal at al. [WHC+06] faced the challenge of needing a system to record audio files for a location-based tour through a historic place and the guides were not able to authentically narrate their stories away from the location. As a solution, they build a mobile application which allows the guides to record their snippets for their tour directly on their PDAs, neglecting audio quality for a more authentic experience [WHC+06].

3.2 Authoring Versus Playing Application

In the examined prototypes, the design of the authoring tools and the normal consumer functionality is approached in one out of two ways. Either, the designers developed two completely separated applications, one for consuming and one for creating location-based content, or they followed the approach of being able to create and change the minimum of each location, but big changes need to happen on a different environment (e.g., on a desktop computer).

An example for the first solution is the game "Premierløytnant Bielke" [WB09]. The authors developed a web interface to easily change the content of each location, allowing non-technical users to maintain and update the content, while the game was implemented as a native mobile application.

The second approach is implemented by Weal et al. [WHC+06]. It is possible to record audio files on the field but heavier changes still need the desktop environment. Another example is the TOTEM application suite [JWBO13]. Jurgelionis et al. created two applications: TOTEM.Designer (desktop application to create data templates) and TOTEM.Scout (mobile application for filling the templates with data in the field).

A third approach would be to enable the mobile application to do all the configuration, creation and maintenance necessary for the tool to work. However, none of the found projects described such a possibility. The described project in this paper follows this approach.

4 Development

Mobilot[2] is a web platform for creating location-based content. Users can create "Mobiduls", modules where they can mark arbitrary locations ("stations") on a map, add media or text and publish such collections. The software has some additional functionality, like collaborative content editing, which supports use cases where multiple users can contribute to a collection of locations, e.g., a collection of all ATMs in a town or a city guide. The application is solely web based, even so the design is highly responsive and optimized for mobile usage. Each feature can easily be accessed on mobile devices. Therefore, in-situ authoring and creation is ensured.

4.1 Technical Stack

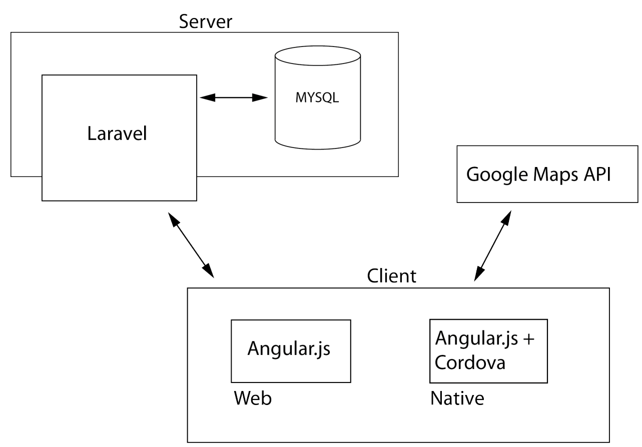


Figure 1: Graphical representation of the architecture.

The architecture of Mobilot can be seen in Figure 1. The back-end (server) is implemented with the PHP MVC framework Laravel, version 4.2, with a MYSQL database. The back-end architecture (server) is mainly a typical REST API with CRUD (Create, Read, Update, Delete) functionality. This kind of architecture is advantageous for creating multiple client applications for different scenarios or devices because the back-end can stay the same.

The front-end (client) is built with the JavaScript framework Angular.js 1.4 . This framework is designed to build big applications as single page applications (SPA). SPAs technically consist of one page even so it

[2]The code base is open source and available at https://github.com/fhstp-mfg/mobilot.

does not seem like it. This has some benefits, like an overall faster experience of the site. Instead of heavy page loads, the applications have to request data only when needed, saving a lot of overhead in HTTP requests. On the downside, there can be a possible slow initial loading of the page when the user visits the page for the first time because the complete application gets loaded at once.

The map feature of Mobilot is realized with the Google Maps API inside the front-end application.

4.2 Goals, Ambitions and Challenges

The existing functionality of Mobilot lets users already easily create location-based content. However, the content was rather static and non-interactive itself. To use the platform in a learning context, we needed some new features to encourage learners to get active. We came up with the idea of extending the existing features to make it possible to easily and fast create scavenger hunt-like learning games. The basic functionality was already in place, the game itself would be a Mobidul created by some kind of game master, like a teacher, and the stations would be the different points of the scavenger hunt. We needed some way to give stations different kind of states (like 'hidden', 'open' or 'completed') and only display those who have already been unlocked.

Another feature, we wanted to implement, was a set of interactive components to enable users to configure the scavenger hunt more freely and create their own rules on how the unlocking of stations would work. These components should be easily configured to the need of the situations.

It was very important to focus on maintaining the mobile and in-situ authoring possibilities of a Mobidul. Therefore, we were challenged not only to create such highly configurable scavenger hunts but also ensuring the mobile use of the editor.

Additionally, we wanted to transform the existing web application into a hybrid version. On the one hand, it will allow us to use native functionality of mobile devices, which could not be used with a web-based application, like access to the Bluetooth chip. On the other hand, we can deploy the application on PlayStore by Google and AppStore by Apple, to be discoverable on those platforms.

4.3 Different States of a Station

The first step to create a scavenger hunt-like experience is to differentiate between various states of a station (see Figure 2). Instead of showing the same information regardless of the context of the user, we introduced four different kinds of conditions with their own rules each station could be in.

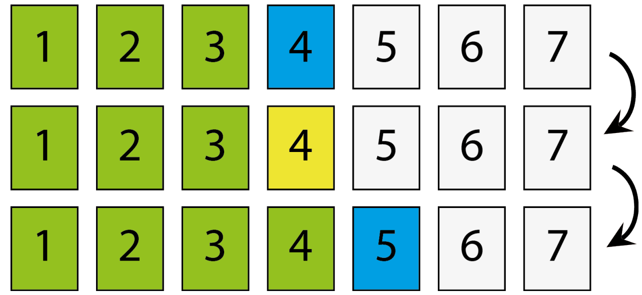


Figure 2: The current (activated) station in the first line is station 4 (blue). All previous stations are completed (green) and all future stations are hidden (white). Then, a player triggers an action to set the current state open (yellow). The station will reveal some new information as well as an interaction component. Finally, once the user has completed the station, the next station will be activated.

Each station starts in hidden state. Hidden stations are not visible on the map or in any list. If players try to access it directly they will be redirected to the currently active station. Once the previous station is completed the station will be activated. Now, the station is visible on the map. In this state players can get challenged to do something in order to proceed with the game, e.g. finding the station. Afterwards, the station can be opened. Technically, the open state is equal to the previous one. The usage of this state is optional but allows new possibilities for game elements. After this, the player will continue with the next station. Finally, a finished station will stay in the completed state. This allows displaying information that should be accessible after playing the game.

4.4 Interaction Components

The scavenger hunt Mobidules offer several interaction elements for configuring a station. Each of these elements offers custom configuration capabilities (e.g., the label on a button or the range of a GPS tracking component) and a call to action in case it gets triggered.

- **HTML Content:** is a very basic non-interactive component for creating simple text and images with a WYSIWYG ("What you see is what you get") editor.
- **Action Button:** is a simple button to trigger an action if clicked. The label and action are configurable.
- **Code Input Field:** is an input field with a submit button. Players will have to enter the correct code in order to trigger the success action. An error action is also implemented.

- **GPS Detector:** If a station is configured with a GPS detector it will check the players' position every five seconds. If the user is within a configurable range an action will be triggered. As a fallback, in case of an inaccurate GPS signal, there will be a code input component to trigger the action anyway.
- **Countdown:** Opening a station with a countdown component will start a timer, that will trigger an action after a specified time.

4.4.1 Actions

To create interactive learning experiences, it is necessary to configure custom responses to user inputs. While some components allow the declaration of two actions, one for success, one for failure, most work with only one of them. The following actions can be triggered:

- **Open this station:** will set the status of the currently active station to "open".
- **Complete this station:** will set the currently active station as completed and the next station as activated. This action can also be used in activated state, which will result in skipping the open state.
- **Say some text:** will open a dialog window with a customizable text. Can be used for hints or custom error messages.
- **Go to current station:** will set the progress to the currently active station. This is useful for completed stations to allow players to quickly navigate back to the currently active station.

4.5 Editor

The design consideration behind the editor was to make it as easy as possible even on mobile devices.

To get to the station editor, one has to either create a new station in the menu or click the edit button in the header at the desired station (see Figure 3). It is required to have editing permits to do so.

The editor is shown in Figure 4. On top, there are four tabs with different options:

- **Base:** allows changing the name of the station as well as its content/configuration.
- **Place:** shows a map for changing the location of the station.
- **Categories:** lets the owner put the station into one or more categories.
- **Options:** allows some additional configuration of the station as well as the option to delete it.

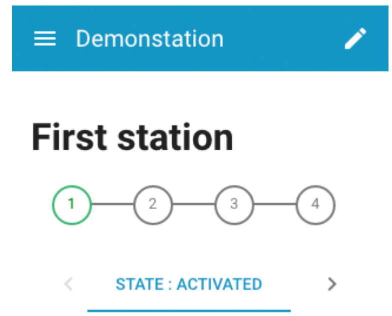


Figure 3: Shows the station view of an Mobidul owner with editing permits. The edit button and the developer tools are visible.

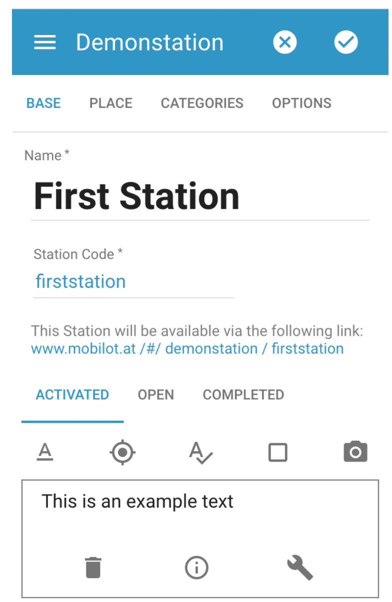


Figure 4: The first view of the station editor. If a new station is created, each state will already contain a HTML element. The user can switch the different states by clicking on the according tab. The currently selected state is indicated by the blue line underneath the name and the different font color.

For creating a scavenger hunt, the first tab (Base) is primarily important because the game logic is managed here. Each state of the station is separated into its own tab. The editor can change between these by touching one of the state names ('activated', 'open' and 'completed').

Underneath the state tabs, the user can add unlimited elements to the configuration by clicking on the desired component, which will add the element to the content area underneath. If an already existing element in the content area is selected it will be displayed with a blue background. The new component is going to be placed directly behind of it.

Each component has an information and delete button. By clicking on the "toolbox" further configuration is possible.

There can only be one opened element to save screen space, which is important to think about when designing for mobile devices. If the user opens a new element the currently selected one will collapse into the preview mode. The order of the elements is changeable with a drag'n'drop gesture. This is consistently used throughout the application to change the order of list items (also used for changing the station order or menu items).

To save the changes, the user has to click on the right button in the header. This will redirect him or her to the station view. The changes can also be discarded by clicking on the cancel button in the left of the save option.

4.5.1 Developer Tools

To make it easier for Mobidul creators to test their content, the application offers developer tools to change their current progress data. These tools can be seen on Figure 3. They consist of two parts. One - the connected dots underneath the station name - shows which station is currently viewed (the outline and font color gets green). It also allows changing stations quickly by clicking one of the other dots. The other feature is right underneath. User can change the state of the station by selecting one of the three tabs. The content will change immediately according to the selection.

5 Evaluation

The goal of the evaluation phase was to test the new game components of the web application. The primary focus was the user experience of the consuming position, not the handling of the content authoring and configuration.

For this evaluation, two different tests were conducted. Both of them were qualitative evaluations with observation followed by group interviews. The target audience of these tests were elementary school children from "Volksschule Tullnerbach" in Tullnerbach, Austria.

Overall, we had 90 participants aged from 6 to 10 (see Table 1 for the distribution of the participants). All participants claimed to have experience in working with mobile devices. Some of the older ones even own smartphones themselves. The tests took place on two days during a project week at the school and were performed on the school grounds. They were instructed

Table 1: Overview of the participant distribution among Mobiduls.

Game	Male	Female	Total
Reading Rally 1	8	15	23
Reading Rally 2	8	10	18
Reading Rally 3	7	2	9
Reading Hike 1	8	14	22
Reading Hike 2	8	10	18
	39	51	90

and supervised by the project leaders and developers of the software. The test content, two separated games with an objective of practicing reading while doing physical activities, was designed by a related teacher. The distinction between the two test modules was based on the age of the participants. Each test took about one hour including instructions at the beginning and a quick group interview at the end.

As test devices we provided iPhone 5, 6 and 6+ as well as iPad 2 and iPad mini.

5.1 Reading Rally (Age 6 to 8)

Around the school, there were five stations at well-known places like the car park or the tennis court. The participants read instructions on their devices to go to a specific station and perform a task. These tasks were different for each station, some were achievable with just using the device, others required extra instruction and validation from a station supervisor (see Figure 5). All tasks ended with receiving a keyword, which had to be entered to complete the station and activate the next one.

Figure 5: Station supervisor instructing the participants.

This test was performed with three classes. Two of the three classes were regular ones with children aged 6 to 8 (with 23 and 18 participants) and one mixed with an age range from 6 to 10 (with 9 participants). Each class got separated into three subgroups and each one of these got three test devices. This meant the children had to share one device with one or two peers. The groups were accompanied by one of the supervisors, in case questions or problems occurred.

The participants were very motivated and rushed from one station to another. However, it was noticeable that the older ones could not immerse as much as their younger peers. Therefore, the mood of the mixed class was a little less enthusiastic. The game was designed for younger children. Therefore, an explanation for this could be that the tasks were too easy and not challenging.

The children were obviously excited about the evaluation devices. They were aware of the cost of the phones and impressed by the responsibility of carrying them. They did not have any problems handling the application, which matches their statement to be experienced in working with smartphones.

Beforehand, it was expected that they would read the instruction in these constellations, but it turned out that one person ended up reading aloud for the whole group. This kid was most times the most confident reader of the class, which made the ambition of the application a bit pointless, because the idea was to practice reading. It even hindered insecure students wanting to hold the device, so they would avoid having to read.

The game design was well received by the participants. Especially, the athletic tasks were highly rated and better memorized than the others during the final interview. The challenge to match numbers with letters was probably too abstract for the target group. From observations it also showed that it was the least engaging station of the rally, mostly leading to having one or two participants solving the task for the whole group. A preference regarding self-instructed stations compared to having a supervisor giving additional information or validation was not stated.

5.2 Reading Hike (Age 9 to 10)

The older participants were tested with another module adapted to their reading abilities. Each one of the two classes got divided into two groups and the children had to share devices in pairs. Each station reveals a new chapter of a continuous story and will give a hint on where to find the next station. Additionally, the current distance to the station is displayed on screen, giving feedback if the moving direction is correct. The application will automatically show the next chapter of the story once the distance was less than $10m$.

The narrative was fitted for the audience. It was the story of a young girl looking for her lost domestic pig called Norbert. While searching him, she meets several other animals and people who help finding her friend. Each one of these encounters is represented by a station and the children walk the same way as the girl in the story.

The game consists of ten stations, the walking distance is about $3km$ and the children needed about one hour to complete it. As a fallback, in case the GPS did not work, there were physical signs attached at each station, providing an unlock code (see Figure 6).

Figure 6: Participants reading the story at a station. The sign with the fallback code is attached to the lamp post.

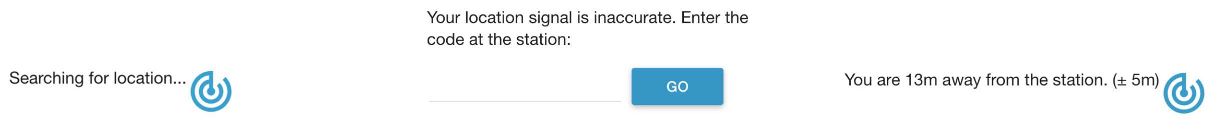
Unlike the reading rally, this game can be played without depending on anything but the smartphone. It should be possible for the participants to come back with their parents and repeat the scavenger hunt with their own devices. Therefore, the interface of the station view is extremely important to be easily understandable without any outside help.

This game got tested with a total of 40 children aged 9 to 10.

During the group interview the students stated that the story was appealing. Some children proposed included more media genres like videos to enhance the experience.

The devices used for this tests were iPad 2, iPad mini and iPhone 6+. The reason for choosing these devices was the assumption that it would be easier for children to read on a larger screen. This decision proved to be false, not only did the iPad suffer the most technical related issues with GPS, but it was also too big and heavy for this kind of usage and target audience.

The main game play element was finding the next station. The most used tool to achieve this objective was the distance meter (see Figure 7). Another feature that would have been useful is the map view, but the participants did not use it at all, which could be the case, because at the initial instructions it was not shown or explained. Wake et al. [WB09] described a similar observation where the participants did not use the map but instead relied on the distance meter very heavily. They try to explain it with the participants - in this case adults - knowing the area very well



(a) (b) (c)

Figure 7: The three states of the GPS component. Loading (a), input field for fallback code (b) and distance meter (c).

and expecting to see different scenarios with younger students. The findings could contradict this thesis, because the participants are within the suggested age range and did not know the area too well.

The participants permanently overrated the accuracy of the distance meter. They tried to find the right direction by walking a few steps and looking at the new distance as feedback. While this seems like a valid approach on solving these problems it implies a few other problems. First of all, the GPS signal is not completely reliable. The game component responsible for calculating the distance was configured to accept an accuracy aviation up to $20m$ before falling back into an alternative mode. The current accuracy was even displayed next to the distance. This effect is also described by Facer et al. [FJS$^+$04] who observed children attributing substantially more intelligence to the devices and technologies than there were in reality [FJS$^+$04].

During the test, a severe bug was noticed the first time. The issue, that occurred when the GPS module of the device returned an error, led to the situation that neither the current position was checked nor the fallback alternative was displayed. To solve this problem, the user had to refresh the page, which triggers a new request for the GPS position, normally resolving the issue.

Another problem was the selection of the fallback codes, which were three random letters. The auto correct feature of the devices made it hard to input these codes and automatically replaced the text with a correction. Words that would pass the auto correct inspection would have been a better choice.

All participants were satisfied with the application. The described problems did not upset them that much. They stated that they would be interested in similar games and could imagine using such an application in their spare time and with their families.

5.3 Summary

Overall, the evaluations, of a self-explanatory Mobidul (Reading Hike) as well as one with additional instructors (Reading Rally), show a successful involvement of the participants in the Mobiduls. Even so the game objectives, like solving a task or progressing in the story, were quite simple, they allowed the children to immerse into the experience. The technical side of the application was stable. However, the GPS availability of some devices had a negative impact on the users' experience. Due to its simplicity, the user interface was easy to use and did not need a lot of instruction.

6 Conclusion

Mobilot shows that it is possible to develop a tool that allows an easy and in-situ creation of location-based learning games. Many games fitting this category show a very similar structure that can be streamlined into a generic system, while maintaining its attractiveness to players.

This paper described the development of a tool to easily create location-based games as well as the outcome of two user tests. The user tests brought good results for the games and the participants enjoyed the game. Still, there is room for improvement. A next step will be the user interface evaluation of the station editor with teachers as the target audience.

The future possibilities of Mobilot seem endless. One of the next extensions will be the implementation of further context recognition like the distance to object (measured with Bluetooth signal) and social interaction. With these new features the tool belt for creating interesting learning experiences will allow the creation of wide-ranging game mechanism. Another idea is to open up the concept of the linear game structure to allow an autonomous exploration of the individual stations or the creation of conditional paths during one game session.

As for further research, it would be interesting to evaluate more mobile learning games in an educational context. Teachers should be encouraged to use technology in their classes and therefore more tools are required to support them.

Acknowledgements

This work was supported by the project seekoi (no. 1154) funded by the Internet Foundation Austria (IPA) as well as the Austrian Ministry for Transport, Innovation and Technology (BMVIT) under the Austrian Security Research Programme KIRAS via the project Courageous Community (no. 850196).

References

[BBS+10] Elizabeth Brown, Dirk Börner, Mike Sharples, Christian Glahn, Tim De Jong, and Marcus Specht. Location-based and contextual mobile learning. A STELLAR Small-Scale Study. STELLAR European Network of Excellence in TEL (EU), 2010. Retrieved 2016-06-29, from `http://oro.open.ac.uk/29886/`.

[CT09] Chih-Ming Chen and Yen-Nung Tsai. Interactive location-based game for supporting effective english learning. In *Proc. of ESIAT*, volume 3, pages 523–526. IEEE, 2009.

[FJS+04] Keri Facer, Richard Joiner, Dana Stanton, Josephine Reid, Richard Hull, and David Kirk. Savannah: mobile gaming and learning? *Journal of Computer assisted learning*, 20(6):399–409, 2004.

[JWBO13] Audrius Jurgelionis, Randall Wetzel, Lauren Blum, and Leif Oppermann. TOTEM. scout: A mobile tool for in-situ creation of location-based content. In *Proc. of IGIC*, pages 89–96. IEEE, 2013.

[KIND+15] Hilda K. Kabali, Matilde M. Irigoyen, Rosemary Nunez-Davis, Jennifer G. Budacki, Sweta H. Mohanty, Kristin P. Leister, and Robert L. Bonner. Exposure and use of mobile media devices by young children. *PEDIATRICS*, 136(6):1044–1050, 2015.

[PP11] Jean-Christophe Puja and David Parsons. A location-based mobile game for business education. In *Proc. of ICALT*, pages 42–44. IEEE, 2011.

[SYOA+14] Christos Sintoris, Nikoleta Yiannoutsou, Alejandro Ortega-Arranz, Rodrigo Lpez-Romero, Menita Masoura, Nikolaos Avouris, and Yannis Dimitriadis. TaggingCreaditor: A tool to create and share content for location-based games for learning. In *Proc. of IMCL*, pages 280–284. IEEE, 2014.

[Tak16] Dean Takahashi. Analytics firms show how pokmon go became a phenomenon, 2016. Retrieved 2016-07-26, from `http://venturebeat.com/2016/07/14/analytics-firms-take-us-inside-the-pokemon-go-phenomenon/`.

[TKK+09] Qing Tan, Kinshuk, Yen-Hung Kuo, Yu-Lin Jeng, Po-Han Wu, Yueh-Min Huang, Tzu-Chien Liu, and Maiga Chang. Location-based adaptive mobile learning research framework and topics. In *Proc. of CSE*, pages 140–147. IEEE, 2009.

[WB09] Jo Dugstad Wake and Rune Baggetun. “premierlytnant bielke”: A mobile game for teaching and learning history. *IJMBL*, 1(4):12–28, 2009.

[WHC+06] Mark J. Weal, Eva Hornecker, Don G. Cruickshank, Danius T. Michaelides, David E. Millard, John Halloran, David C. De Roure, and Geraldine Fitzpatrick. Requirements for in-situ authoring of location based experiences. In *Proc. of MobileHCI*, pages 121–128. ACM, 2006.

Klangsalat – Auditives Navigationssystem

Andreas Negrei, BSc
dm151555@fhstp.ac.at

Florian Csizmazia, BSc
dm151504@fhstp.ac.at

Tamás Künsztler, BSc
dm151553@fhstp.ac.at

Mathias Berger, BSc
dm151502@fhstp.ac.at

Patrick Wiertel, BSc
dm151539@fhstp.ac.at

Florian Pichler, BSc
dm151559@fhstp.ac.at

Maximilian Sramek, BSc (Hons)
dm151532@fhstp.ac.at

Abstract

Das Ziel dieser Arbeit war die Implementierung einer Android-Applikation mit binauralen Klangtexturen, die den NutzerInnen eine Stadtführung durch St. Pölten auf auditiver Ebene ermöglicht. In der App sind 22 Points of Interest (POIs) auf einer Google Maps Karte definiert, denen jeweils ein spezifischer Klang zugeordnet ist. Diese Klangtexturen sind akustische Simulationen der jeweiligen Orte. Die Schallquellen sind, wie deren zugehöriger POI, statisch in der App positioniert und besitzen einen klar definierten Klangradius. Bewegt sich also ein User innerhalb des Klangradius einer Schallquelle, so beginnt diese zu klingen. Jede Klangtextur eines POI wiederholt sich und wird je nach Entfernung vom NutzerIn zum POI intensiver bzw. weniger intensiv und binaural aus der richtigen Richtung wahrgenommen.

1 Einleitung

Die verbreitete Nutzung von Smartphones und immer genauere GPS- und Richtungssensoren bieten EntwicklerInnen zahlreiche Möglichkeiten zur Implementierung standortbasierter Anwendungen. Das Smartphone ist als täglicher Begleiter oft auch Ersatz für klassische Straßenkarten. Während die zielorientierte Verwendung von Google Maps und Co. vergleichsweise wenig Erlebnisnutzen mit sich bringt, finden auch standortbasierte Spiele, wie beispielsweise Pokémon Go, immer mehr Gefallen.

Ziel dieser Arbeit war die Entwicklung einer standort- und richtungsbasierten App (in weiterer Folge als Klangsalat bezeichnet), mit der Erweiterung um die akustische Ebene. Signifikante Orte der Stadt St. Pölten, Points of Interest (POIs) genannt, werden durch bestimmte Sounds repräsentiert. Die Entfernung der NutzerInnen zu einem POI bestimmt die Lautstärke des jeweiligen Sounds. Durch gezieltes Drehen des Körpers (bzw. des Smartphones) und Bewegung der NutzerInnen kann die St. Pöltner Innenstadt, mit Hilfe von visueller und akustischer Navigation, erkundet werden.

Eine wichtige Rolle spielt der Einsatz von Binauralität (siehe Abschnitt 4.1), auch als Richtungshören bezeichnet. Bei der Bewegung des Smartphones und der damit einhergehenden Änderung der Winkel zu den einzelnen POIs, sollten die Sounds weiterhin aus der standortspezifisch richtigen Richtung zu hören sein. Diese Positionierung auf der horizontalen Ebene ist mit einfacher Stereophonität nicht bzw. in keinem zufriedenstellenden Ausmaß zu erreichen.

Am Ende des ersten Projektabschnittes sollte ein Android-Prototyp der App Klangsalat stehen, der die akustische Ortung der NutzerInnen auf die Probe stellt. Obwohl die visuelle Ebene in Form einer Karte trotzdem bedient wird, stellt sich auch die Frage, ob die rein akustische Navigation durch die zu den POIs gehörigen Sounds möglich und sinnvoll ist. Die Auswahl und Implementierung der richtigen Sounds stellte einen entsprechend wichtigen Teil der Entwicklungsarbeit dar. Im Rahmen einer Fokusgruppe wurden daher die POIs und mögliche Sounds gemeinsam mit potenziellen NutzerInnen diskutiert. Usertests sollten die Funktionalität der App allgemein, die richtige Auswahl der Sounds und die korrekte Richtungswahrnehmung evaluieren.

In: W. Aigner, G. Schmiedl, K. Blumenstein, M. Zeppelzauer (eds.): Proceedings of the 9th Forum Media Technology 2016, St. Pölten, Austria, 24-11-2016, published at http://ceur-ws.org

2 State of the Art

Das Projekt „Klänge der Stadt – Soundscape Regensburg“ (List, Mantaj, & Ottmann, 2014) demonstrierte bereits einen akustischen Stadtplan mit Hilfe einer Google Maps Karte. Hierbei wurden für bestimmte Orte die Originalklänge zugeordnet und in der Soundmap mit dazugehörigem Foto markiert – jedoch ohne mobile Applikation. Eine solche wurde als Prototyp für Tablets von Georg Weidenauer im Rahmen seiner Diplomarbeit „Virtual Soundscape Elements“ (Weidenauer, 2014) mit Hilfe von „KEMAR Impuls Antworten“ implementiert – inklusive Einbindung virtueller Schallquellen. Eine App für die Aufzeichnung von Soundscapes ist auch beim Projekt „Record the Earth“ (Pijanowski, 2014) in Verwendung. Die dabei verwendete Methode nennt sich „Soundscape Ecology“ und dient, ähnlich wie bei dem hier vorgestellten Projekt, zur Kategorisierung von Klängen. Durch die daraus gewonnenen Informationen und Eigenschaften von Sounds wurde beim Projekt „Euronoise“ (Davies et al., 2009) die Tiefenstaffelung gebildet. Zusätzlich wurden Aspekte der Sounddefiniton vom Experiment „The measurment of soundscapes – is it standardizable?“ (Genuit & Fiebig, 2014) aufgegriffen, um die emotionalen Reaktionen von NutzerInnen zu evaluieren. Mit den Attributen von Soundscapes beschäftigten sich Ljungdahl Eriksson und Berg bereits 2009 in ihrer Arbeit „Soundscape Attribute Identification“ (Ljungdahl Eriksson & Berg, 2009), 2014 der Holländer Almo Farina in „Soundscape Ecology“ (Farina, 2014), sowie die 2015 veröffentlichte Arbeit „Effects of Psychoacoustical Factors on the Perception of Musical Signals in the Context of Environmental Soundscape“ (Deng, Kang & Liu, 2015).

Die Verbindung einer Applikation mit GPS Daten behandelte das bereits erwähnte Projekt „Record the Earth“ – jedoch ausschließlich zur Lokalisation von aufgenommenen Sounds. Die Annäherung, Entfernungsunterschiede, sowie die Verwendung von einem binauralen System war Teil des Projekts „Interactive 3-D Audio“ (Schmidt, Schwartz, & Larsen, 2012) und „I Hear NY3D“ (Musick, Andreopoulou, Boren, Mohanraj, & Roginska, 2013). In der aktuellen Smartphone-Applikation „Pokémon Go!“ ist die erfolgreiche Echtzeit-Interaktion von App- und GPS-Daten – also das Zusammenspiel zwischen Funktionen und Standort der UserIn – in Verwendung. „Project GO!“ (“Project GO! Trailer,” n.d.) arbeitet innerhalb einer App mit aktuellem und geschichtlichem Bildmaterial, welches positionsbezogen abgerufen wird. Ausgehend vom dargestellten Stand der Forschung wurde mit Klangsalat eine App entwickelt, wo einzelne Klangtexturen zum Teil aus Klängen der Vergangenheit und aus der Gegenwart gebildet werden.

3 Hard- und Software Spezifikationen zur Umsetzung

Für die Umsetzung dieses Projekts wurde eine mobile Android App entwickelt. Programmiert wurde in der IDE (Integrated Development Environment) „Android Studio“ mit der Programmiersprache Java. Die Signalverarbeitung erfolgt mit einem Pure Data Patch. Pure Data[1] (Pd) ist eine Open Source[2] und visuelle Programmiersprache, die v.a. in der Audio Branche oft zum Einsatz kommt. Für den Einsatz von Pd Patches auf Smartphones sind nur Knoten (Nodes), die im Paket Vanilla Pd enthalten sind, verfügbar. Weiters muss eine spezielle Library (libpd)[3] eingebunden werden, welche die Kommunikation zwischen dem Client Code (Java) und dem Pure Data Patch ermöglicht. Da im Rahmen einiger Tests auf verschiedenen Mobiltelefonen unterschiedliche Winkeldaten ausgegeben wurden, beschränkte man sich in der Alpha Version der Applikation auf das Mobiltelefon Sony Xperia Z3. Es verfügt über einen 2,5 GHz Quad-Core-Prozessor und 3 GB RAM. Das Gerät besitzt außerdem einen internen Kompass, Gyrometer und GPS Sensoren. Um auf Sensor- und GPS-Daten zugreifen zu können, muss die Google Play Services API mit eingebunden werden.

Zur Verarbeitung der POIs wurden diese mit Namen, Beschreibung und Koordinaten in einer CSV-Datei gespeichert, die als Liste in Java geladen und verarbeitet werden kann. Diese Liste ist einfach wart- und erweiterbar und benötigt nur wenig Speicherplatz.

Obwohl für optimales Richtungshören In-Ear Kopfhörer verwendet werden sollten, wurde bei den Tests – vor allem aus hygienischen Gründen – auf halb-offene, aufliegende Kopfhörer zurückgegriffen.

4 Methoden

Zur Findung der passenden Klänge für die einzelnen POIs wurde eine Diskussionsrunde mit Personen unterschiedlichen Alters, verschiedenen beruflichen Hintergründen und unterschiedlichen Wohnorten durchgeführt. Diese Fokusgruppe bestand aus elf TeilnehmerInnen, denen zu jedem Point of Interest ein Bild gezeigt wurde, zu dem sie jeweils die Klänge beschreiben sollten, die ihnen dazu einfallen. Die Moderatoren gaben zu den POIs – wenn nötig – kurze Erklärungen. Durch die Diskussion sollten zu jedem

[1] https://puredata.info

[2] Open Source bezeichnet freie Software, deren Quellcode frei verfüg- und erweiterbar ist

[3] https://puredata.info/downloads/libpd

POI verschiedene Klänge und Assoziationen gefunden werden, gleichzeitig aber auch ein Konsens zum intuitiv "richtigen" Klang. Zu jedem Point of Interest wurden nach Übereinstimmung der Probanden zumindest vier verschiedene Klänge vorgeschlagen. Aus der Liste der Ergebnisse wurden im Anschluss vom Projektteam jene Sounds, die der Klangästhetik der zu produzierenden Klangtexturen am besten dienten, ausgewählt. Für die Produktion der Klangtexturen wurden einerseits Klänge in St. Pölten mittels eines Aufnahmegerätes mit Stereomikrofonierung (XY-Mikrofon des Zoom H6) und einem Richtrohrmikrofon aufgenommen, andererseits wurde auf lizenzfreie Sounds der Soundbibliothek www.freesound.org zurückgegriffen. Dabei wurden nur Sounds verwendet, die unter der CC0 (creative commons 0 – Public Domain Dedication) lizensiert sind. Diese Lizenz erlaubt die Kopie, Veränderung und Verbreitung der Klänge (auch für kommerzielle Zwecke). Die durchschnittliche Länge einer Klangtextur beträgt 30 Sekunden. Da diese beim Benutzen der Applikation so lange wiederholt werden, bis sich der User nicht mehr in deren Klangradius befindet, war es wichtig, dass keine Übergangsartefakte zwischen End- und Anfangspunkt der Klangtexturen wahrnehmbar sind. Die Klangtexturen wurden daher so produziert, dass eine beliebige Wiederholung ohne eindeutige Identifizerung von Start und Ende möglich ist. Die in St. Pölten aufgenommenen Klänge und jene von www.freesound.org wurden in die Software Ableton Live importiert und dort editiert. Jede Klangtextur besteht aus mehreren Klangebenen, die zeitlich und klanglich (Filterung, Pitch-Shifting, Time-Stretching, Kompression, Hall) aufeinander angepasst wurden.

4.1 Binauralität

Das binaurale Hören ermöglicht ein Richtungshören beziehungsweise ein räumliches Hören. Laut Blauert und Braasch werden Schallsignale, die auf den Kopf treffen, bezüglich ihres Frequenzspektrums linear verzerrt. Es kommt zu Änderungen von Amplituden- und Phasenverlauf des Spektrums. Diese Änderungen, die auf dem Übertragungsweg zwischen Schallquelle und Ohren passieren, können mit Hilfe von sogenannten Head-Related Transfer Functions (HRTFs) mathematisch beschrieben und messtechnisch erfasst werden. Hierfür werden aus dem Nahfeld und in bestimmten Winkeln rund um den Kopf Impulse abgespielt und mittels Mikrofonen im Ohr der Versuchsperson oder eines Kunstkopfes aufgenommen. Durch Laufzeit- und Form-Unterschiede ergeben sich spezielle Filterkurven, die beim Hören einen sehr realistischen, räumlichen (binauralen) Richtungseindruck geben. Der Binauraleffekt in der Stereofonie ist mit Kopfhörern am deutlichsten wahrzunehmen. Das liegt daran, dass beim Abhören mit Kophörern die Signale am linken und am rechten Ohr völlig getrennt voneinander ankommen. (Jens Blauert & Jonas Braasch, 2008, Kapitel 3: Räumliches Hören, Herausgeber: Stefan Weinzierl, S. 89-90)

Bei der entwickelten App „Klangsalat" wird Binauralität vor allem für das Orten der Richtung, aus der die Klänge der Points of Interests kommen, verwendet. Die Applikation berechnet den Winkel zwischen den GPS Koordinaten der NutzerIn und des POIs in Bezug auf den magnetischen Nordpol. Das Ergebnis dieser Berechnung wird vom aktuellen Blickwinkel der NutzerIn (abhängig von der Ausrichtung des Smartphones), der über die Sensordaten erhalten wird, abgezogen.

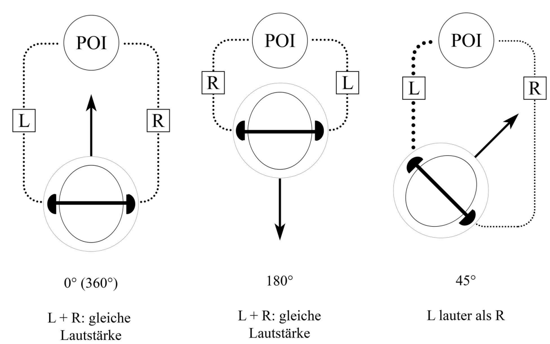


Abbildung 1: Änderung der Lautstärke in Bezug auf den Blickwinkel

Wenn das Ergebnis 0 oder 360 beträgt, blickt der/die NutzerIn direkt auf den Point of Interest. Beträgt das Ergebnis 180, hat der/die BenutzerIn dem POI den Rücken zugedrehte.

Durch die Drehung des des Smartphones ändert sich der erreichnete Winkel und auch die wahrgenommene Lautstärke. Die Tatsache, dass sich die Klangtextur bei der Drehung nicht mitbewegt führt dazu, dass der POI akustisch geortet werden kann.

5 Implementierung

Die Programmierung von Klangsalat wurde in zwei Bereiche unterteilt. Die Java-seitige Entwicklung in Android Studio, umfasste die Berechnung der Entfernungsdaten zu den POIs und die Verarbeitung der Sensor- und GPS-Daten. In Pure Data wurden Binauralität, Lautstärkeanpassung und das Abspielen der Sounds in Pd umgesetzt.

Das eingebaute GPS-Modul des Mobiltelefons liefert eine auf ein bis drei Meter genaue Standortlokalisation mit Werten für Breiten- und Längengrad. Aus dem aktuellen GPS-Standort und den Koordinaten der POIs, die aus der CSV–Datei geladen werden, wird die Distanz zu jedem POI berechnet und die nähesten vier werden mit folgenden Parametern an Pd gesendet:

- ID (zur Auswahl der richtigen Audiodatei)
- Entfernung (für die Lautstärkeanpassung)
- Winkel (für HRTFs und Binauralität)

Im Folgenden ist ein Ausschnitt aus jener Funktion zu sehen, die bei jedem Standort-Update (ca. alle 20 Sekunden) durch das Smartphone aufgerufen wird, um die notwendigen Parameter für die vier nächsten POIs an Pd zu senden.

```
// send the 4 pois to pd
for (int i = 0; i < poisToSend.size(); i++) {
    // get current poi
    Poi currentPoi = poisToSend.get(i);

    if(currentPoi != null) { // double check that current poi is not null
        // send relevant values distance, angle and id
        // to the corresponding pd receives
        PdBase.sendFloat("distance" + (i+1), currentPoi.getDistance());
        PdBase.sendFloat("angle" + (i+1), currentPoi.getAngle());
        // check if current poi is already playing
        // to avoid new sound initialisation
        if(!currentPoi.playing)
            PdBase.sendFloat("id" + (i+1), currentPoi.getId());
    }
}
```

Abbildung 2: Senden der 3 Parameter zu Pd

Davor erfolgt eine Überprüfung, ob sich der POI innerhalb eines Radius von 200 Metern ausgehende vom Standort der NutzerIn befindet. Die Entfernungsberechnung zwischen zwei Punkten ist durch folgende Formel gegeben:

$$a = sin^2\left(\frac{\Delta\varphi}{2}\right) + cos\varphi_1 * cos\varphi_2 * sin^2\left(\frac{\Delta\varphi}{2}\right)$$

$$c = 2 * atan2(\sqrt{a} * \sqrt{(1-a)})$$

$$d = R * c$$

φ ... *Latitude*

λ ... *Longitude*

R ... *Erdradius (6,371 km)*

Formel 1: Berechnung des Klangradius

Mithilfe von Gyrometer und Kompass des Smartphones erfolgt die Winkelberechnung, welche für die Richtungsgebung der verschiedenen Sounds herangezogen wird.

Die Einbindung von libpd in Android ermöglicht die Kommunikation zwischen der App und dem Pure Data Patch. Die Klasse PDBase erlaubt den Zugriff auf Methoden, die diese Kommunikation ermöglichen. Um die oben genannten Parameter an Pd zu schicken wird die Methode *sendFloat* verwendet (Brinkmann, 2012).

Die übermittelten Winkeldaten werden zur Bearbeitung im Pd Patch für die binaurale Verteilung der Sounds herangezogen. Es werden Filterkurven für beide Ohren aus den Impulsantworten (IR) der Binaural Library[4] “1002” des IRCAM Instituts in Tables (Subpatches, die Arrays und die dazugehörigen Graphen beinhalten) im 15-Grad-Abstand gespeichert. Diese werden abhängig vom aktuellen Winkel über die jeweiligen Klangtexturen gefalten, um eine Lokalisation der Soundquelle zu ermöglichen. Die relativ grobe Auflösung von 15 Grad ist für die Applikation dennoch ausreichend, wie die Usertests (siehe Abschnitt 6) beweisen. Zusätzlich wird mit zunehmender Entfernung (0-200 Meter) zum jeweiligen POI die Lautstärke des Sounds abgeschwächt.

$$M = \frac{L * 0{,}4 + 200}{1{,}617}$$

Formel 2: Abschwächung der Lautstärke einzelner POI's

L nimmt dabei Werte zwischen 0 und 200 Metern ein. Der Mulitplikator M wird zum Schluss mit einem "dbtorms" Befehl umgerechnet, um so einen ansprechenden Lautstärkenverlauf zu erhalten, der Werte zwischen 0 und 8 einnehmen kann. Die Werte wurden im Laufe der Entwicklung und des Testens ermittelt und angepasst.

Nach Anpassung aller Parameter wird der Befehl zum Abspielen gegeben und an den Audioausgang des Mobiltelefons geschickt.

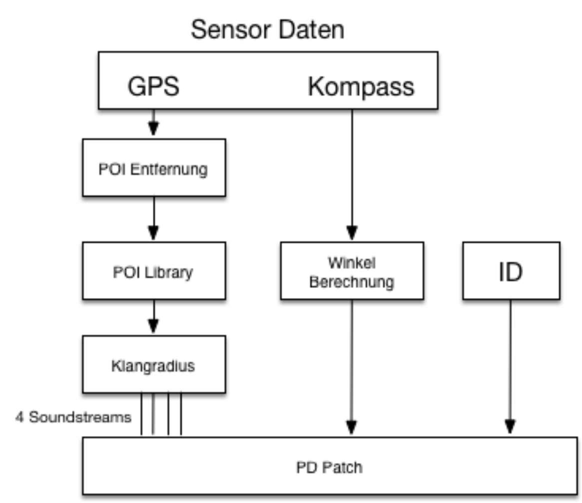


Abbildung 3: Ablauf der Berechnungen in Java und die Kommunikation zum Pd Patch

[4]http://recherche.ircam.fr/equipes/salles/listen/download.html

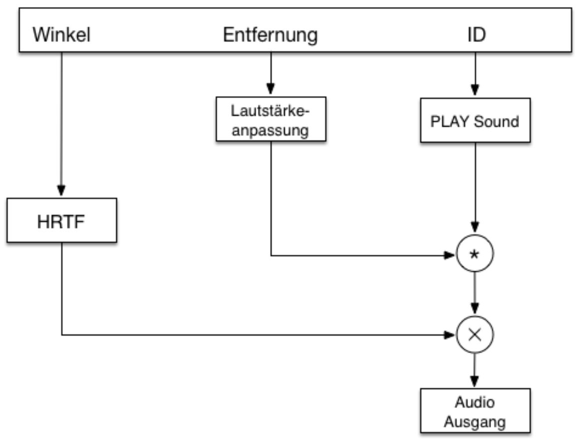


Abbildung 4: Ablauf der Berechnungen im Pd Patch

6 Empirische Studien/Evaluierung

Die richtige Funktion der Applikation wurde in mehreren Schritten beurteilt. Im Laufe der Implementierung wurden mehrere Testläufe durchgeführt, bevor schlussendlich mit dem fertigen Prototyp mit projektfremden Personen ein Usability-Test im finalen Zielgebiet vorgenommen wurde.

6.1 Testläufe

In den Testläufen nach den einzelnen Implementierungsphasen wurde die richtige Funktion der jeweiligen Elemente getestet. Dazu gehörten zum Beispiel das richtige Erkennen der Winkel und Abstände vom Gerät zu den einzelnen POIs, die richtige Ausgabe der Werte zur Weiterverarbeitung in Pure Data und die Einbindung des Pure Data Patches in die Applikation.

Auf der Seite der Klangverarbeitung in Pd mussten unter anderem die Änderungen des adaptierten Patches, das flüssige Funktionieren der Binauralität und die Entfernungswahrnehmung, die Einbindung der produzierten Sounds in die Applikation und die Verschachtelung mehrerer Instanzen getestet werden.

6.2 Rundgang St. Pölten

Nach dem ersten Zusammenfügen aller Einzelteile und dem Übertragen auf zwei Testgeräte führte das Projektteam einen Rundgang im finalen Zielgebiet durch. Dabei sollten grundsätzliche Probleme der Applikation identifiziert und anschließend behoben werden, bevor diese zur Evaluierung an Dritte weitergereicht werden konnte.

Zu den Problemen gehörten zum Beispiel die Vertauschung von links und rechts, Probleme beim Abspielen einzelner Sounds, kein automatisches Hinzukommen bzw. Wegfallen der Instanzen und digitale Verzerrungen, wenn Sounds lauter wurden. Außerdem wurde die Qualität und Funktionalität der einzelnen Sounds beurteilt, wie zum Beispiel Verständlichkeit oder Identifizierbarkeit.

6.3 Rundgang Wien

Vor dem Usability-Test wurden die durchgeführten Adaptierungen noch einmal live getestet. Dieser weitere Rundgang fand in einer Fußgängerzone in Wien statt. Die ursprünglichen St. Pöltner POIs wurden dabei gleichmäßig im Testgebiet verteilt. Teile der in Abschnitt 6.2 beschriebenen Probleme wurden durch Anpassung der gesendeten Paramter an Pd bzw. durch eine Umpolung bei der Auswahl der Impulsantworten behoben. Um die entstandenen Verzerrungen zu entfernen wurden die Sounds für die Endversion noch einmal normalisiert.

6.4 Usability-Test

Der Usability-Test wurde wieder im finalen Zielgebiet durchgeführt. Zwei Teams waren mit jeweils einem Gerät und Kopfhörern im Kerngebiet St. Pölten unterwegs und baten Passanten um ihre Mithilfe. Nach einem Rundgang von maximal 5-10 Minuten sollte ein kurzes Feedback in Form eines Fragebogens abgegeben werden. Dabei sollten Grundprinzip und Funktion der Applikation durch Fragen, die von „trifft sehr zu" bis „trifft gar nicht zu" beantwortet werden konnten, bewertet werden. Dies sollte von einer möglichst diversen und realitätsnahen Zielgruppe passieren.

Von den Testpersonen kamen 50 % nicht aus St. Pölten, die anderen 50 % waren entweder in der Stadt ortskundig oder wohnhaft. Zwei Drittel waren unter 26 Jahre alt und 58,3 % weiblich.

Für 75 % der ProbandenInnen war die Funktion der Applikation klar verständlich. Nur 8,3 % gaben an, dass das eher nicht zutrifft. 83,3 % gaben an, dass die Sounds sehr oder eher zu den festgelegten POIs passen. Für drei Viertel der Personen war die Orientierung mithilfe der Klangtexturen sehr oder eher klar. Von den Befragten fühlte sich die Hälfte eher nicht und 8,3 % gar nicht durch das Benutzen der Applikation von ihrer Umgebung abgelenkt, 41,7 % hingegen sehr oder eher schon. Bei der Frage, ob die angezeigte Karte der Umgebung für die Orientierung notwendig sei, waren die Antworten eher ausgeglichen, mit Tendenz ins Positive: 25 % finden ja, 33,3 % eher ja, 25 % eher nein und 16,7 % gar nicht.

- Es ist für mich verständlich, wie die Applikation funktioniert. (12 Antworten)

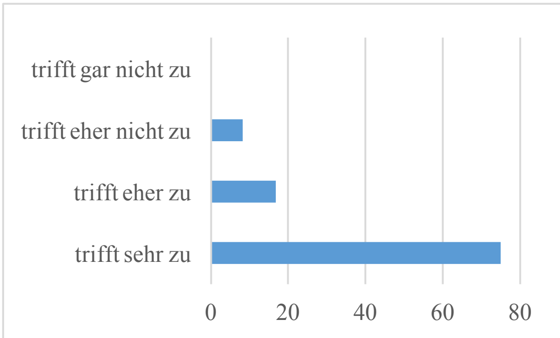


Abbildung 5: Verständlichkeit der App

- Die Orientierung ist klar und ich weiß, in welche Richtung ich laut Applikation gehen muss. (12 Antworten)

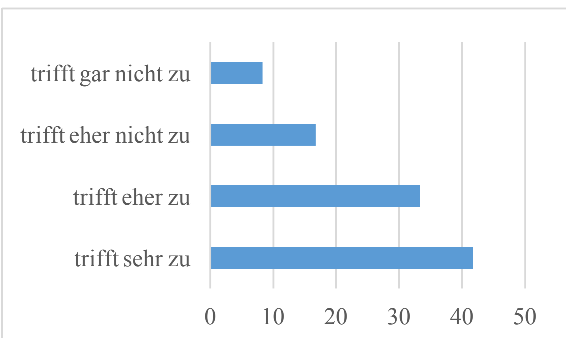


Abbildung 6: Orientierung

- Ich werde von der Applikation sehr abgelenkt und mir fällt es schwer mich auf die Umgebung zu konzentrieren. (12 Antworten)

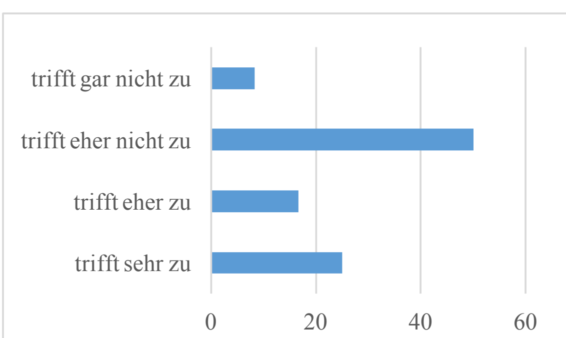


Abbildung 7: Ablenkung und Sicherheit

- Die Sound sind zu den jeweiligen Orten passend gewählt. (12 Antworten)

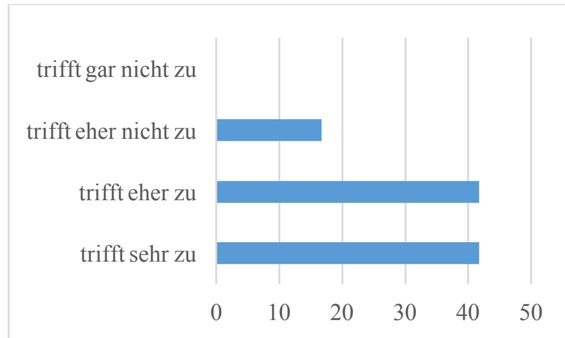


Abbildung 8: Qualität der Sounds

Da nicht alle ProbandInnen alle POIs und deren Sounds gehört haben, wurden die Routen zusätzlich mit auf den Fragebögen vermerkt, um eine individuelle Bewertung der jeweiligen Sounds zu ermöglichen.

7 Diskussion

7.1 Findung der richtigen Sounds

Ziel des Projektes war die Entwicklung eines Android-Prototypen der Applikation mit Sounds zu den enthaltenen 22 Points of Interest (POIs). Zur initialen Findung der Sounds wurde, wie in Abschnitt 4 beschrieben, eine Diskussion mit einer Fokusgruppe durchgeführt. Bei dieser Diskussion zeigte sich auch deutlich, wie viele verschiedene Sounds die TeilnehmerInnen mit den jeweiligen POIs assoziieren. Während in der initialen Idee noch von Einzelsounds ausgegangen wurde, wurden auf Basis all dieser Ideen für viele POIs stattdessen Soundcollagen entwickelt. Durch die Verwendung verschiedener Sounds in einer Collage sollte für möglichst viele NutzerInnen eine eindeutige Zuordenbarkeit zu den jeweiligen POIs gegeben sein.

Bei den Usertests in St. Pölten zeigte sich, dass die Auswahl der Sounds zu den jeweiligen POIs den Erwartungen der NutzerInnen entsprach. Es gab wenig Probleme mit der Vermischung der Sounds unterschiedlicher POIs. Die Tatsache, dass für 83 % der TeilnehmerInnen die Aussage „Die Sounds sind zu den jeweiligen Orten passend gewählt.“ eher oder sehr zutrifft, unterstreicht, dass die richtigen Sounds gefunden und entwickelt wurden. Auch dort, wo mehrere POIs in einem sehr kleinen Radius liegen (so zum Beispiel am St. Pöltner Rathausplatz), konnten die NutzerInnen durch gezieltes Drehen des Kopfes (bzw. des Smartphones) oder durch Bewegung in eine bestimmte Richtung die POIs und ihre Sounds separieren.

Die binaurale Darstellung half also bei der Orientierung und Lokalisierung der POIs.

Neben den Ergebnissen des Fragebogens wurden auch im persönlichen Gespräch während des Usertests Fragen beantwortet und Erkenntnisse dokumentiert. Hier zeigte sich, dass auch Probleme bei der Unterscheidbarkeit der POIs und ihrer Sounds aufgetreten sind. Als Beispiel sind hier vor allem die POIs „Sparkasse“ und „Marktviertel“ zu nennen. In beiden Soundcollagen kommt das Klirren von aufeinanderprallenden Münzen vor (siehe auch Anhang 1). Mehrere TeilnehmerInnen des Usertest empfanden die Ähnlichkeit als verwirrend oder sogar störend. Es zeigte sich, dass die Unterscheidbarkeit der einzelnen Sounds wichtig für die richtige Orientierung ist.

7.2 Richtige Funktion des Prototyps

Der bereits angesprochene Usertest hatte neben der Evaluation der erstellten Sounds und ihrer richtigen Zuordenbarkeit auch den Test der richtigen Funktion der Applikation zum Ziel. Durch die Beschränkung der Benutzeroberfläche auf ein Minimum sollte den NutzerInnen die Möglichkeit gegeben werden, sich voll und ganz auf die Funktionsweise der Sounds zu konzentrieren. Die persönlichen Gespräche während des Usertests bestätigten diese Annahme.

Grundsätzlich war die Funktion des Prototyps in diesem Bereich gegeben und zufriedenstellend. Mit der Bewegung der UserInnen wurden die Lautstärken der jeweils spielenden Sounds angepasst, die Drehung des Smartphones führte zur Anpassung der Filterkurven und damit zur binaural richtigen Darstellung der Sounds im Raum. Probleme gab es allerdings bei sehr geringer Entfernung der NutzerInnen zu einem POI. Die Maximallautstärke war zu hoch und das Drehen des Smartphones führte zu einem „Springen“ des Sounds im Raum. Hierbei muss über eine Anpassung der Filterkurven bei geringer Distanz oder über eine Abschwächung der durch den Sensor erhaltenen Winkeldaten nachgedacht werden.

Außerdem wurden Probleme mit der Performance der App auf den verwendeten Smartphones festgestellt. Obwohl beim Test Smartphones der Oberklasse mit schnellen Prozessoren und ausreichend Arbeitsspeicher zum Einsatz kamen (Sony Xperia Z3 und OnePlus Two), konnte die App nur nach Beenden aller anderen Prozesse zufriedenstellend verwendet werden. Bei Sounds mit hohem Signallevel (z.B. beim Sound „Festspielhaus“) wurden außerdem Probleme wie Rauschen und Knacksen festgestellt.

8 Ausblick

Bei der Entwicklung des Prototyps wurde vor allem auf die grundsätzliche Funktion der Applikation und den Einsatz der richtigen Sounds Wert gelegt. Entwickelt wurde ausschließlich für Android auf Basis des Testgeräts Sony Xperia Z3. Grundfunktionen, wie das Zulassen der Standortüberprüfung durch die UserInnen, sind zwar enthalten, allerdings noch nicht ausreichend getestet und auf Fehler überprüft. Weiters muss bei einer Weiterentwicklung großes Augenmerk auf die Performance der App gelegt werden. Häufiges Abstürzen oder Probleme mit der Signalverarbeitung und daraus resultierendes Rauschen oder Knacksen müssen analysiert und behoben werden. Weitere Punkte sind das Branding (zum Beispiel bei der Darstellung der Karte) und Entwicklung bzw. Design der Benutzeroberfläche. Dazu gehört auch die Darstellung von Informationen zu den jeweiligen POIs, die für die NutzerInnen nach der akustischen Navigation zu einem POI einen Zusatznutzen darstellen soll.

Nach Umsetzung der besprochenen Punkte kann außerdem über eine Entwicklung bzw. Anpassung der App für iOS-Geräte nachgedacht werden. Die kommerzielle Verwertbarkeit der Applikation an sich ist als eher gering anzusehen, sie soll nach der Fertigstellung aber gratis im Google Play Store herunterladbar sein.

Das System – Location basierte und binaurale Sounds auf Smartphones – kann allerdings auch über die Grenzen der Applikation Klangsalat hinaus interessant sein. So ist es zum Beispiel in den immer beliebter werdenden Standort basierten Smartphone-Spielen, wie Pokémon Go, denkbar, binaurale Sounds einzusetzen und die Spiele so um die akustische Ebene zu erweitern. Klangsalat kann hier als Beispiel für die Verwendung von akustischer Orientierung in Smartphone-Apps dienen.

Literaturverzeichnis

Brinkmann, P. (2012). *Making musical apps* (1st ed). Sebastopol, Calif: O'Reilly Media.

Davies, W. J., Adams, M. D., Bruce, N. S., Cain, R., Jennings, P., Carlyle, A., … Plack, C. (2009). Euronoise 2009.

Deng, Z., Kang, J., & Liu, A. (2015). Effects of Psychoacoustical Factors on the Perception of Musical Signals in the Context of Environmental Soundscape. In *Audio Engineering Society Convention 138*. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=17701

Farina, A. (2014). *Soundscape Ecology*. Dordrecht: Springer Netherlands. Retrieved from http://link.springer.com/10.1007/978-94-007-7374-5

Genuit, K., & Fiebig, A. (2014). The measurment of soundscapes - Is it standardizable?

List, V., Mantaj, S., & Ottmann, S. (2014). Soundscape Regensburg. Retrieved July 29, 2016, from https://soundscaperegensburg.wordpress.com/

Ljungdahl Eriksson, M., & Berg, J. (2009). Soundscape Attribute Identification. In *Audio Engineering Society Convention 126*. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=14989

Musick, M., Andreopoulou, A., Boren, B., Mohanraj, H., & Roginska, A. (2013). I Hear NY3D: An Ambisonic Installation Reproducing NYC Soundscapes. In *Audio Engineering Society Convention 135*. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=16997

Pijanowski, B. C. (2014). Record the Earth. Retrieved July 29, 2016, from https://www.recordtheearth.org/

Project GO! Trailer. (n.d.). Retrieved September 10, 2016, from https://vimeo.com/118826123

Schmidt, M., Schwartz, S., & Larsen, J. (2012). Interactive 3-D Audio: Enhancing Awareness of Details in Immersive Soundscapes? In *Audio Engineering Society Convention 133*. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=16522

Weidenauer, G. (2014). *Virtual Soundscape Elements* (Diplomarbeit). FH St.Pölten, Etsdorf am Kamp.

Weinzierl, S., & Verband Deutscher Tonmeister (Eds.). (2008). *Handbuch der Audiotechnik*. Berlin: Springer.

Anhang 1 Abgebildete POIs und deren Sounds

ID	Name	Sound
1	Rathausplatz	Stimmengewirr, Gespräche, Straßenmusiker (Ziehharmonika)
2	Rathausgasse 2	Schubertiade, Klavier
3	Bahnhof St. Pölten	Zug, Gleisgeräusche, Stimmen, Durchsage
4	Klosterviertel	Geistlicher Männergesang in Kirche
5	Holzviertel	Sägen, Klopfen, Hämmern, Schleifen, Holzarbeiten
6	Ledererviertel	Kuh, Leder ziehen, Lederproduktion
7	Marktviertel	Einzelne Münze fällt, Hund bellt, Kassieren, Holzwagen fährt
8	Zentrale städtische Feuerwehr	Brennen, Knacken, Feuerwehrhorn, Doppler-Effekt, Löschen
9	Synagoge	Kantor-Gesang/-Gebet
10	Riemerplatz	Völkl-Marsch, Relief
11	Domplatz	Ausgrabungen, düstere, flächige Musik, Kirchenglocken
12	Franziskaner Kirche	Brunnen-Rauschen, Türe öffnet und schließt, Chorgesang in Kirche
13	Domgasse 4	Akustische Solo-Gitarre
14	Sparkasse	Geldautomat, Münzen klirren
15	Regierungsviertel	Telefone, Büro, Papierrascheln, Stimmen, geschäftiges Treiben
16	Klangturm	Stille (Hintergrund: der Klangturm klingt nicht mehr)

17	Festspielhaus	Orchester stimmt, Klatschen, Stimmen
18	ORF Landesstudio NÖ	Radio rauschen, alten Fernseher einschalten, Fragmente Musik aus Radio
19	Schützenkompanie	Marschieren, Schüsse, Blasmusik
20	Hammerpark	Bachrauschen, Vogelgezwitscher
21	Traisen	Flussrauschen, Fahrrad auf Schotter
22	Kaserne	Flugzeugmotoren, entfernte Kanonen, Blasmusik, Marschieren

Convenient Mobile Usability Reporting with UseApp

Johannes Feiner
IIT[1]
johannes.feiner@fh-joanneum.at

Keith Andrews
IICM[2]
kandrews@iicm.edu

Elmar Krainz
IIT[1]
elmar.krainz@fh-joanneum.at

IICM[2], Institute for Information Systems and Computer Media (IICM), Graz University of Technology, Austria
IIT[1], Institute for Internet Technologies & Applications (IIT), FH JOANNEUM, Austria

Abstract

Usability reporting is necessary to communicate the results of usability tests to developers and managers. Writing usability reports (data aggregation, interpretation, formatting, and writing for specific readerships) can be tedious. Particularly for mobile usability evaluation, where recording user task performance outside a lab is often necessary, testing and reporting can be costly. In many cases, automated extraction of usability findings would be helpful, but is rather difficult to achieve with commonly used report formats such as Word or PDF.

UseApp is a tablet-based web application developed to overcome some of these limitations. It supports the capture of usability data *in the field* during testing, simplifying data collection and aggregation. Live-reports are generated on-the-fly and usability findings can be exported electronically to bug tracking systems.

1 Mobile Usability Reporting

Usability evaluations are performed to validate the usability (and user experience) of software products. For example, experts might conduct heuristic evaluations (HE) to detect potential flaws in applications based on their experience and judgement. Thinking aloud (TA) tests might be conducted with representative test users to discover problems in realistic usage scenarios.

Smaller software development teams often do not have the resources to conduct extensive user studies. Furthermore, the modern practice of agile software development encourages rapid, incremental testing. In both cases, testing has to be simple and efficient. A tool supporting electronic capture of usability data as easily as using a pen and paper can be of great benefit.

In: W. Aigner, G. Schmiedl, K. Blumenstein, M. Zeppelzauer (eds.): Proceedings of the 9th Forum Media Technology 2016, St. Pölten, Austria, 24-11-2016, published at http://ceur-ws.org

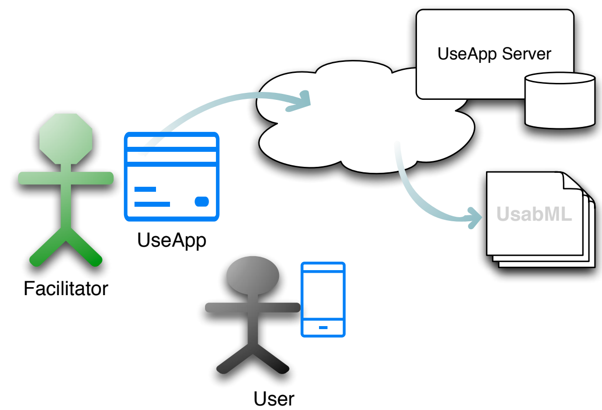


Figure 1: UseApp helps evaluators record the performance of users electronically during usability testing.

Nowadays, many applications are designed to run on mobile phones, shifting the focus of usability testing to mobile usability testing. This shift requires a tool set supporting mobile reporting. Reports in structured formats such as UsabML [FAK10] allow evaluation results to be processed electronically: findings can be extracted and then imported into bug tracking systems automatically.

2 Related Work

Many methods for evaluating user interfaces have been developed over the past three decades [DR93; Nie95]. Formative usability evaluation [Red+02] seeks to discover potential usability problems during the development of an interface, so they can be fixed. Of the formative evaluation techniques, Heuristic Evaluation (HE) [NM90; Nie94b; HLL07] and Thinking Aloud (TA) [RC08] testing are particularly widely used. Nielsen [Nie94a] suggested that some, simplified usability evaluation is always better than none. Brooke [Bro96] proposed the System Usability Scale (SUS) to make assessment through questionnaires simpler and results comparable through normalised scoring.

Usability reporting feeds back the findings of usability evaluations to development teams [Que05; Her16]. According [Yus15] and [YGV15] using conventional bug tracking systems for usability reporting does now work well. Structured written reports have traditionally been used [FH03; LCA97]. Some efforts have been made to standardise the structure of such reports, including the Common Industry Format (CIF) [NIS99] for formal experiments. However, usability reports are still largely delivered in traditional document formats such as Microsoft Word and PDF, which are extremely hard to process automatically. UsabML [FAK10] is a structured, XML-based format format for usability reports, which allows tool-based extraction and/or conversion and thus fosters simpler automation and reuse.

Reporting usability defects to software developers (cmp [Hel+11]) is a challenge still. [YGV16] investigated reporting and analysed 147 responses. They detected a gap between the what reporters provide and what developers need when fixing defects. UseApp aims into the same direction, as it narrow this gap by supporting semiautomated handover of usability results into bug tracking systems.

Modern usability evaluation shifted towards open use situations and takes the mobile context into account, as discussed in [BH11], [KSV12] and [Lan13]. ISO standards support objective measurement of the usability of mobile applications as reported in [MIA16]. Several tools to assist mobile usability testing can be found in literature. [Sto+15] present MARS, a mobile app rating scale. The tool helps assessing and classifying apps in health sector. The challenges of automatic UI observation and event logging to improve usability on mobile apps can be found [Ma+13], but the support for usability engineering methods (like TA or HE) is missing. Frameworks with a set of different tools and methods to support mobile usability evaluation can be found at [And+01] and [Che16].

Also, some commercial products are on the market. For example, Usertesting[1] is a product which helps to add usability testing on mobile platforms. Beside premium/paid support for testing, simple test can be created with the help of an online tool. Another tool for testing-support of mobile web sites is UXRecorder[2] which supports recording of users touch and facial impressions.

The systematic mapping study [ZSG16] about mobile application testing techniques categorised the different approaches and stated that 19 out of 79 studies employed usability testing. The paper discusses many challenges of mobile testing, such as context-awareness, lab vs. in-the-wild testing, video recording or mobile eye-tracking. One of the main challenges addressed in several papers was *Improving the test suite*. Furthermore, [ZSG16] refer to research groups working on improved toolkits and testing frameworks: [Can+13] for Advanced Test Environment (ATE), a platform which supports automatic execution of user experience tests, [LH12] for a toolkit for unsupervised evaluation of mobile applications, [BH09] a logging based framework to evaluate usability of apps on mobile devices, and [VCD15] for automated mobile testing as a service. For research crowdsourcing in mobile testing [Sta13] created the lightweight *Cloud Testing of Mobile Systems* (CTOMS) framework.

In contrast to this work, which focuses on reporting, only few of the cited approaches mention post-processing and reuse of reports at all.

Criteria	Description	Usage in UseApp
Simple and Fast	Minimise input, use templates.	No pen and paper required. Placeholders and default values.
Context Awareness	Sensor support (GPS), timing.	Auto-timing of task duration.
Don't Repeat Yourself (DRY)	Manage and store project and user details.	Reuse existing user details, questionnaires.
Export and Reuse	Structured formats, post-processing.	Export as UsabML.

Table 1: Selected design criteria for a mobile usability reporting tool.

3 UseApp Concept

UseApp is a client-server web application, as shown in Figure 1. The facilitator manages the mobile user testing and typically enters data into the system using a web browser on a tablet. The criteria used to design UseApp are shown in Table 1. Data entry should be fast and simple, through a minimal interface and use of templates, placeholders, and default values. Sensors should be used to automate procedures as far as possible. Data should only have to be entered once. Overviews and reports should be generated on-the-fly and results should be exported in a structured reporting format.

Recipes for common evaluation tasks, such as a thinking aloud test or administering a standard questionnaire should be available pre-canned. The interface should support focus and context: giving an overview whilst simultaneously allowing the facilitator to focus on the details of current actions. Colour-coded indicators should give feedback about already completed sections, and highlight where data is still

[1] https://www.usertesting.com/.
[2] http://www.uxrecorder.com/.

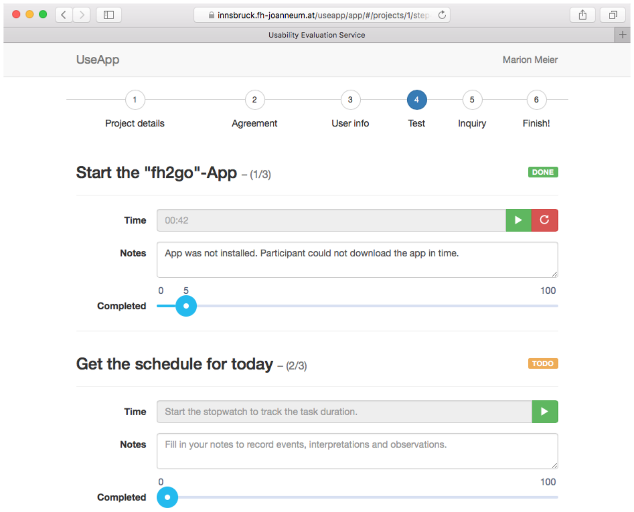


Figure 2: Just six steps – from setup, to data entry to a final report – motivates even small development teams to perform mobile usability tests.

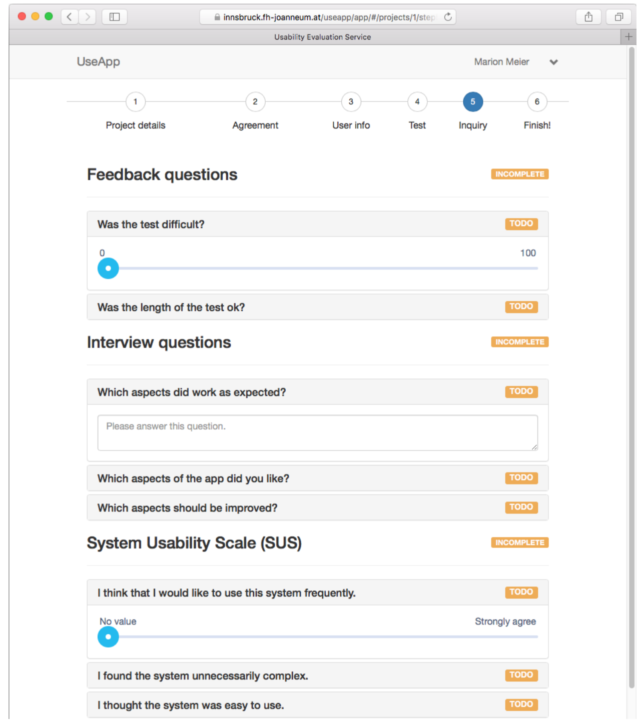


Figure 3: Many built-in placeholders and the timer functionality allow simple and fast reporting.

incomplete. To remove the need for pen and paper, everything should be possible directly on the tablet: from signing a consent form with a stylus or via audio, to answering questionnaire questions by tapping.

4 UseApp Implementation

UseApp currently has built-in support for running Thinking Aloud (TA) tests and administering SUS [Bro96] questionnaires. In future versions, support for Heuristic Evaluations (HE) and other questionnaires and rating scales will be added.

The client implementation uses many features of modern HTML5 web technologies, in order to support the features outlines in Section 3. Responsive web design is used to support several screen resolutions and provide sensible fallbacks where features are not supported by a particular device or browser. Offline storage, sensors, audio input and output, and canvas-based charts are all used.

The UseApp server is built in Ruby/Rails and exposes a restful application programming interface (API). Thus, the client only retrieves and stores data on the server, but the layout and rendering are completely server independent.

The workflow for a TA test comprises six steps (Project Details, Agreement, User Info, Test, Inquiry, and Finish), as indicated in the top bar in Figure 2. The workflow starts with entering the project details. Test users then give their consent and answer demographic and background questions.

The facilitator can track individual or collective performance directly with help of *UseApp*. Placeholders and templates support and speed up facilitator input as shown in Figure 3. Timing of task duration is supported by built-in timers. Task completeness can be indicated just be moving a slider. After completing tasks, the users are asked for feedback. As the questions have all been prepared and assembled in advance, the answers are collected in electronic form.

The results can be viewed for single participants, or for a group of participants, including means and summaries. Multiple charts are available to support interpretation and communication of the results. Figure 4 shows an example. Notes and annotations can be added by the facilitator.

5 UseApp in Action

UseApp was trialled for a number of mobile usability evaluations. The UseApp server was set up in-house and the installation of the web app on an iPad was prepared in advance. The manager of each study entered the project details, task descriptions, and questionnaire questions in advance. As users performed their tasks, the facilitator had their iPad in hand to guide the session, enter observations, and record task duration. After completing the tasks, an interview was conducted and a questionnaire was filled out. Immediately after each test user has finished, the usability managers had access to the results and could add any comments or notes relevant to that test.

Feedback from the first users of UseApp (the usability evaluations managers and facilitators) has indicated some of its benefits and limitations:

- *Feedback*: the top bar indicating the six steps to completion was useful feedback.

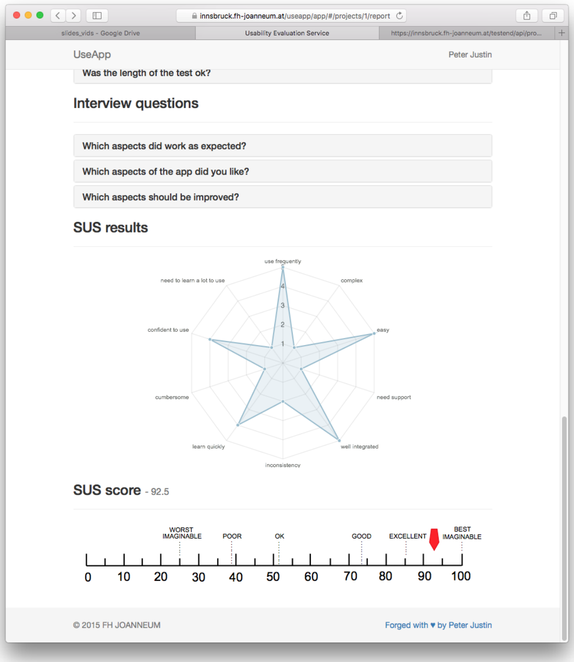


Figure 4: Results are calculated on-the-fly and make results available instantly.

- *No Paper*: no need for paper keeps the test environment uncluttered.
- *Re-Use*: where user testing is required multiple times, the reuse of already prepared evaluation documents, such as same or similar questions for the questionnaire, is time saving.
- *Export*: software developers liked the idea of post-processing reports. After exporting the usability reports in structured UsabML, automated import into bug tracking systems is not difficult.

UseApp acts as a practical companion when running a mobile usability test. Although UseApp can help, preparing and conducting usability tests still takes time and effort.

A minor limitation was the lack of support for freehand writing when signing the consent form. A tablet supporting a stylus might be useful for future versions instead of forcing users to draw with their fingers.

6 Concluding Remarks

UseApp has the potential to support usability evaluators in multiple ways. It simplifies data entry when conducting mobile usability tests, provides templates for input, automation for recording tasks, and reuse of project data. Instant reporting and flexible export into structured UsabML help accelerate the provision of usability findings by the test team to the appropriate software developers.

Ongoing improvement of UseApp will expand evaluation methods supported and the palette of built-in templates. The use of GPS sensors to track location might also be useful in some evaluation contexts.

References

[And+01] Terence S. Andre, H. Rex Hartson, Steven M. Belz, and Faith A. McCreary. “The User Action Framework: A Reliable Foundation For Usability Engineering Support Tools”. In: *International Journal of Human-Computer Studies* 54.1 (Jan. 2001), pages 107–136. doi:10.1006/ijhc.2000.0441 (cited on page 2).

[BH09] Florence Balagtas-Fernandez and Heinrich Hussmann. “A Methodology and Framework to Simplify Usability Analysis of Mobile Applications”. In: *Proc. IEEE/ACM International Conference on Automated Software Engineering*. ASE 2009. IEEE Computer Society, 2009, pages 520–524. ISBN 9780769538914. doi:10.1109/ASE.2009.12 (cited on page 2).

[BH11] Javier A. Bargas-Avila and Kasper Hornbæk. “Old Wine in New Bottles or Novel Challenges: A Critical Analysis of Empirical Studies of User Experience”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’11. Vancouver, BC, Canada: ACM, May 7, 2011, pages 2689–2698. ISBN 9781450302289. doi:10.1145/1978942.1979336 (cited on page 2).

[Bro96] J. Brooke. “SUS: A Quick and Dirty Usability Scale”. In: *Usability Evaluation in Industry*. Edited by Patrick W. Jordan, Bruce Thomas, Bernard A. Weerdmeester, and Ian L. McClelland. Taylor & Francis, 1996. Chapter 21, pages 189–194. ISBN 0748404600 (cited on pages 2–3).

[Can+13] Gerado Canfora, Francesco Mercaldo, Corrado Aaron Visaggio, Mauro DAngelo, Antonio Furno, and Carminantonio Manganelli. “A Case Study of Automating User Experience-Oriented Performance Testing on Smartphones”. In: *Proc. 6th International Conference on Software Testing, Verification and Validation*. ICST 2013. Mar. 18, 2013, pages 66–69. doi:10.1109/ICST.2013.16 (cited on page 2).

[Che16] Lin Chou Cheng. “The Mobile App Usability Inspection (MAUi) Framework as a Guide for Minimal Viable Product (MVP) Testing in Lean Development Cycle”. In: *Proc. 2nd International Conference in HCI and UX on*

Indonesia 2016. CHIuXiD 2016. Jakarta, Indonesia, Apr. 13, 2016, pages 1–11. ISBN 9781450340441. doi:/10.1145/2898459.2898460 (cited on page 2).

[DR93] Joseph S. Dumas and Janice Redish. *A Practical Guide to Usability Testing*. Ablex, Dec. 1993. ISBN 089391990X (cited on page 1).

[FAK10] Johannes Feiner, Keith Andrews, and Elmar Krajnc. "UsabML - The Usability Report Markup Language: Formalising the Exchange of Usability Findings". In: *Proc. 2nd ACM SIGCHI Symposium on Engineering Interactive Computing Systems*. EICS 2010. Berlin, Germany: ACM, May 2010, pages 297–302. ISBN 1450300839. doi:10.1145/1822018.1822065 (cited on pages 1–2).

[FH03] Andy P. Field and Graham Hole. *How to Design and Report Experiments*. Sage Publications, Feb. 2003. ISBN 0761973826 (cited on page 2).

[Hel+11] Florian Heller, Leonhard Lichtschlag, Moritz Wittenhagen, Thorsten Karrer, and Jan Borchers. "Me Hates This: Exploring Different Levels of User Feedback for (Usability) Bug Reporting". In: *Proc. Extended Abstracts on Human Factors in Computing Systems*. CHI EA 2011. Vancouver, BC, Canada: ACM, May 7, 2011, pages 1357–1362. ISBN 9781450302685. doi:10.1145/1979742.1979774 (cited on page 2).

[Her16] Morten Hertzum. "A Usability Test is Not an Interview". In: *interactions* 23.2 (Feb. 2016), pages 82–84. doi:10.1145/2875462 (cited on page 2).

[HLL07] Ebba Thora Hvannberg, Effie Lai-Chong Law, and Marta Kristín Lérusdóttir. "Heuristic evaluation: Comparing ways of finding and reporting usability problems". In: *Interacting with Computers* 19.2 (2007), pages 225–240. doi:10.1016/j.intcom.2006.10.001. `http://kth.diva-portal.org/smash/get/diva2:527483/FULLTEXT01` (cited on page 1).

[KSV12] Artur H. Kronbauer, Celso A. S. Santos, and Vaninha Vieira. "Smartphone Applications Usability Evaluation: A Hybrid Model and Its Implementation". In: *Proc. 4th International Conference on Human-Centered Software Engineering*. HCSE 2012. Toulouse, France: Springer-Verlag, Oct. 29, 2012, pages 146–163. ISBN 9783642343469. doi:10.1007/978-3-642-34347-6_9. `http://dx.doi.org/10.1007/978-3-642-34347-6_9` (cited on page 2).

[Lan13] Tania Lang. "Eight Lessons in Mobile Usability Testing". In: *UX Magazine* 998 (Apr. 10, 2013). `https://uxmag.com/articles/eight-lessons-in-mobile-usability-testing` (cited on page 2).

[LCA97] Darryn Lavery, Gilbert Cockton, and Malcom P. Atkinson. "Comparison of Evaluation Methods Using Structured Usability Problem Reports". In: *Behaviour & Information Technology* 16.4 (1997), pages 246–266. doi:10.1080/014492997119824 (cited on page 2).

[LH12] Florian Lettner and Clemens Holzmann. "Automated and Unsupervised User Interaction Logging As Basis for Usability Evaluation of Mobile Applications". In: *Proc. 10th International Conference on Advances in Mobile Computing & Multimedia*. MoMM 2012. Bali, Indonesia: ACM, 2012, pages 118–127. ISBN 9781450313070. doi:10.1145/2428955.2428983 (cited on page 2).

[Ma+13] Xiaoxiao Ma, Bo Yan, Guanling Chen, Chunhui Zhang, Ke Huang, Jill Drury, and Linzhang Wang. "Design and implementation of a toolkit for usability testing of mobile apps". In: *Mobile Networks and Applications* 18.1 (2013), pages 81–97 (cited on page 2).

[MIA16] Karima Moumane, Ali Idri, and Alain Abran. "Usability Evaluation of Mobile Applications Using ISO 9241 and ISO 25062 Standards". In: *SpringerPlus* 5.1 (2016), page 1. doi:10.1186/s40064-016-2171-z (cited on page 2).

[Nie94a] Jakob Nielsen. *Guerrilla HCI – Using Discount Usability Engineering to Penetrate the Intimidation Barrier*. Jan. 1994. `http://www.nngroup.com/articles/guerrilla-hci/` (cited on page 1).

[Nie94b] Jakob Nielsen. *Ten Usability Heuristics*. Nielsen Norman Group. 1994. `https://www.nngroup.com/articles/ten-usability-heuristics/` (cited on page 1).

[Nie95] Jakob Nielsen. "Usability Inspection Methods". In: *Conference Companion on Human Factors in Computing Systems*. CHI '95. Denver, Colorado, United States: ACM, 1995, pages 377–378. ISBN 0897917553. doi:10.1145/223355.223730 (cited on page 1).

[NM90] Jakob Nielsen and Rolf Molich. "Heuristic Evaluation of User Interfaces". In: *Proc. Conference on Human Factors in Computing Systems*. CHI '90. Seattle, Washington, USA: ACM, 1990, pages 249–256. ISBN 0201509326. doi:10.1145/97243.97281 (cited on page 1).

[NIS99] NIST. *Common Industry Format for Usability Test Reports*. National Institute of Standards and Technology. Oct. 1999. http://zing.ncsl.nist.gov/iusr/documents/cifv1.1b.htm (cited on page 2).

[Que05] Whitney Quesenbery. *Reporting Usability Results – Creating Effective Communication*. Tutorial Slides. Dec. 2005. http://www.wqusability.com/handouts/reporting_usability.pdf (cited on page 2).

[Red+02] Janice (Ginny) Redish, Randolph G. Bias, Robert Bailey, Rolf Molich, Joe Dumas, and Jared M. Spool. "Usability in Practice: Formative Usability Evaluations — Evolution and Revolution". In: *Extended Abstracts on Human Factors in Computing Systems*. CHI EA 2002. Minneapolis, Minnesota, USA: ACM, 2002, pages 885–890. ISBN 1581134541. doi: 10.1145/506443.506647 (cited on page 1).

[RC08] Jeffrey B. Rubin and Dana Chisnell. *Handbook of Usability Testing: Howto Plan, Design, and Conduct Effective Tests*. 2nd edition. John Wiley & Sons, May 2008. ISBN 0470185481 (cited on page 1).

[Sta13] Oleksii Starov. "Cloud Platform For Research Crowdsourcing in Mobile Testing". Master's Thesis. East Carolina University, Jan. 2013. http://thescholarship.ecu.edu/bitstream/handle/10342/1757/Starov_ecu_0600M_10953.pdf (cited on page 2).

[Sto+15] Stoyan R. Stoyanov, Leanne Hides, David J. Kavanagh, Oksana Zelenko, Dian Tjondronegoro, and Madhavan Mani. "Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps". In: *JMIR mHealth and uHealth* 3.1 (2015), e27. doi:doi:10.2196/mhealth.3422. http://mhealth.jmir.org/2015/1/e27/ (cited on page 2).

[VCD15] Isabel Karina Villanes, Erick Alexandre Bezerra Costa, and Arilo Claudio Dias-Neto. "Automated Mobile Testing as a Service (AM-TaaS)". In: *Proc. IEEE World Congress on Services*. June 2015, pages 79–86. doi:10.1109/SERVICES.2015.20 (cited on page 2).

[Yus15] Nor Shahida Mohamad Yusop. "Understanding Usability Defect Reporting in Software Defect Repositories". In: *Proc. 24th Australasian Software Engineering Conference*. ASWEC ' 15 Vol. II. Adelaide, SA, Australia: ACM, 2015, pages 134–137. ISBN 9781450337960. doi:10.1145/2811681.2817757 (cited on page 2).

[YGV15] Nor Shahida Mohamad Yusop, John Grundy, and Rajesh Vasa. "Reporting Usability Defects: Limitations of Open Source Defect Repositories and Suggestions for Improvement". In: *Proc 24th Australasian Software Engineering Conference*. ASWEC ' 15 Vol. II. Adelaide, SA, Australia: ACM, Sept. 28, 2015, pages 38–43. ISBN 9781450337960. doi:10.1145/2811681.2811689 (cited on page 2).

[YGV16] Nor Shahida Mohamad Yusop, John Grundy, and Rajesh Vasa. "Reporting Usability Defects: Do Reporters Report What Software Developers Need?" In: *Proc. 20th International Conference on Evaluation and Assessment in Software Engineering*. EASE '16. Limerick, Ireland: ACM, May 2016, 38:1–38:10. ISBN 9781450336918. doi:10.1145/2915970.2915995 (cited on page 2).

[ZSG16] Samer Zein, Norsaremah Salleh, and John Grundy. "A Systematic Mapping Study of Mobile Application Testing Techniques". In: *Journal of Systems and Software* 117 (2016), pages 334–356. http://www.ict.swin.edu.au/personal/jgrundy/papers/jss2016.pdf (cited on page 2).

Performance Comparison between Unity and D3.js for Cross-Platform Visualization on Mobile Devices

Lorenz Kromer, Markus Wagner, Kerstin Blumenstein,
Alexander Rind, Wolfgang Aigner

St. Poelten University of Applied Sciences, Austria

Abstract

Modern data visualizations are developed as interactive and intuitive graphic applications. In the development process, programmers basically pursue the same goal: *creating an application with a great performance.* Such applications have to display information at its best way in every possible situation. In this paper, we present a performance comparison on mobile devices between D3.js and Unity based on a Baby Name Explorer example. The results of the performance analysis demonstrated that Unity and D3.js are great tools for information visualization. While Unity convinced by its performance results according to our test criteria, currently Unity does not provide a visualization library.

1 Introduction & Related Work

Visualization systems provide interactive, visual representations of data [CMS99] designed to help people understand complex phenomena and augment their decision-making capabilities [Mun14]. Given the interconnectedness of the current age and the increasing volumes of collected data, there is a dire need for such support. While many usage scenarios can be identified in scientific research and business management, systems for personal visualization [HTA$^+$15] and casual information visualization [PSM07] serve exceptionally broad audiences. These visualizations focus less on task-driven activities and more on curiosity and enjoyment while exploring personally relevant data. Showing trends of popular baby names, the Name Voyager [Wat05] is a typical example of a casual visualization.

A main challenge faced by the developers of casual visualization systems is the heterogeneity of devices and platforms they should support. In particular for the casual context, mobile phones and tablets are more suitable than classical desktop computers [BWA15a, BNW$^+$16, HTA$^+$15, LAMR14]. Native systems, e.g., apps for Android or Apple, are only runnable on the platform for which the code is compiled for. Cross-platform support requires the development on top of different software stacks and to maintain separate code bases. One approach to address this challenge are web-based visualizations, i.e. using web technology such as D3.js [BOH11] within the browser. However, a wide-spread concern is that web-based systems lack performance. For example, Baur stated in a 2013 interview [BSB13] that for big visualization systems such as TouchWave [BLC12], going native cannot be avoided because “in the web it looks like a slide show”. Besides the negative effects of interactive latency [LH14], performance overheads negatively affect battery load of mobile devices. Alternative approaches are cross-compilers such as Unity [uni16], which can deploy a single code base to native systems for multiple platforms. Yet, a limitation of Unity is that it does not include a software library for visualizing data [WBR$^+$16]. These two approaches for cross-platform visualization work very differently during both implementation and runtime. The choice will largely depend on the respective application scenario but empirical data on their performance is needed to inform such a decision.

While some research has been carried out to compare the performance of different web-based visualization technologies [LAB$^+$08, JJK08, KSC12], no studies have been found which compare the performance of web-based and cross-compiled visualization approach.

In: W. Aigner, G. Schmiedl, K. Blumenstein, M. Zeppelzauer (eds.): Proceedings of the 9th Forum Media Technology 2016, St. Pölten, Austria, 24-11-2016, published at http://ceur-ws.org

(a) Unity

(b) D3.js

Figure 1: Shows a screenshot of the Baby Name Explorer interface implemented with (a) Unity and (b) D3.js. Shows the circle packing chart (left) with the corresponding grouped bar chart (right) representing the frequency per year for male (blue) and female (pink) names.

Neither could we identify performance results obtained from different target platforms.

Thus, the paper at hand, contributes a performance comparison between Unity (cross-compiled to native) and D3.js (web-based) on four mobile devices. For this, we created two implementations of a casual visualization system to explore popular baby names as described in Section 2. Section 3 covers the implementation details and test setup. After the test results in Section 4, we conclude our work in Section 5 and outline future work.

2 Visualization Design

As proof of concept we started with implementing a simple interactive visualization setup using an open data set of the regional government of Upper Austria on the 50 most often used male and female baby names from 2004 to 2013. The dataset includes the variables *name* (nominal), *gender* (categorical), *year* (quantitative) and *count* (quantitative). All these data are merged together into a table provided as *.csv file. As visualization concepts we combined a circle packing chart [HBO10] with grouped bar charts [CM84].

Initially, the circle packing chart shows the first letters of the baby names as bubbles and its diameter matches with the number of babies per year. A slider is positioned at the bottom of the screen for selecting the year to display.

By tapping a bubble, the bubble expands and the names which are related to the first letter are shown inside the big bubble (see Figure 1). The color of a name bubble is related to the gender (pink := female, blue := male) and the diameter matches the number of babies with the name for the selected year. During the layout phase, the bubbles are placed using physics-based movement like gravity and the biggest bubble is set to the center of the screen. The circle packing chart is linked with a grouped bar chart. The bar chart initially shows the number of babies for all names grouped per year, split into female and male names (using the same colors as for the bubbles). When selecting a first letter bubble, the grouped bar chart shows the number of babies for names starting with the selected letter. When selecting a name bubble (e.g., “Leonie”), the grouped bar chart changes to a single bar chart presenting the number for the name per year.

3 Implementation and Test Setup

To introduce the implementation and test setup, we describe the used tools for implementation D3.js and Unity, the four test devices and environments, the performance criteria and desired results as well as the measured values and methods.

3.1 Test Devices and Environments

Since we focus on cross-platform visualization, the test devices cover a range from tablets (Nexus 9 and iPad Air) to Smartphones (iPhone 6S+ and Galaxy S6 Edge). Both visualization systems are investigated on the devices shown in Table 1.

When selecting the mobile test devices, we deliberately choose devices with larger screen sizes, since the presentation of the tested visualization (see Section 2) on a screen size of 5” or small is not optimal.

The visualization is tested under Android 5.1 (Nexus 9 and Galaxy S6 Edge) and iOS 9 (iPad Air and iPhone 6S+). In addition to the requirements of the devices, the test concept of this paper also examines the dependencies of both visualization versions of external components such as libraries and plug-ins, which were used during the development process.

Unity: With the development environment of

Table 1: Overview of the dimensions of the test devices.

Device	**Type**	**Screen size**	**Screen resolution**	**Processor**	**RAM**	**Graphics processor**
Nexus 9	Tablet	8.9”	$2048 \times 1536px$	NVIDIA Tegra K1	2 GB	NVIDIA GeForce ULP
iPad Air LTE	Tablet	9.7”	$2048 \times 1536px$	Apple A7	1 GB	PowerVR G6430
iPhone 6S+	Smart-phone	5.5”	$1920 \times 1080px$	Apple A9	2 GB	PowerVR GT7600
Galaxy S6 Edge	Smart-phone	5.1”	$2560 \times 1440px$	Samsung Exynos 7 Octa 7420	3 GB	Mali-T760 MP8

Unity it is possible to make a project accessible for multiple platforms. The Unity version of the Baby Name Explorer (Figure 1a) is exported in two versions (Android and iOS). The rich development environment of the game engine Unity includes a sufficient repertoire of physics components and 3D elements. Therefore, we did not have to use external libraries.

D3.js: Since the implementation of the visualization in D3.js (Figure 1b) is web browser based, we used the Google Chrome web browser as test environment which is available on all tested devices (see Table 1). Thus, the visualization is represented under the same technological conditions. For the implementation of the web based version, we did not need additional JavaScript libraries, because D3.js contains all functionalities.

3.2 Measured Values and Methods

To compare a number of software applications, common metrics and measurement points have to be defined [MFB+07]. Subsequently the used methods are:

- **FPS:** For measuring the frames per second (FPS) rates, time logging functions are added around rendering methods in the code, logging the results via logfiles or the console.
- **CPU utilization:** To show the difference between the hardware components, the CPU utilization was observed while performing both visualizations in a specific scenario and five minutes in idle mode. Therefore, it was ensured that no other processes were running on the device.
- **Loading time of raw data:** Both version (Unity and D3.js) contain an explicit function to load the raw data. In order to compare the raw data loading from a CSV file, the elapsed time was measured between the explicit function call and end.

In relation to the technical implementation, Unity and D3.js are strongly different. To overcome this issue, we recorded the system parameters and console logs with OS specific development systems, because there are no uniform functions available to detect the previously listed system parameters.

With the aforementioned measured values, both visualization systems were tested in a specific user scenario. In this case, the Baby Name Explorers usage was simulated over 60 seconds by a regular interaction with the respective system. To reduce the effects of operating system and other processes beyond user control, this user scenario was repeated five times on each visualization system per tested device.

4 Results

The results of the performance comparison of both versions are separated into the three measured parameters, which were presented before. All the measured values of the different test devices were compared into an Excel sheet for preprocessing. By using MS Excel, we processed the calculation of the median values to eliminate outliers and exported the result for each parameter as grouped bar chart.

4.1 CPU Usage Analysis

Based on the performed measurements, Unity generates less CPU usage than D3.js. Calculating the median across all measured devices, Unity takes 22% and D3.js takes 38%. Figure 2 illustrates a diagram to compare the CPU usage between the tested devices in idle mode and while performing both versions.

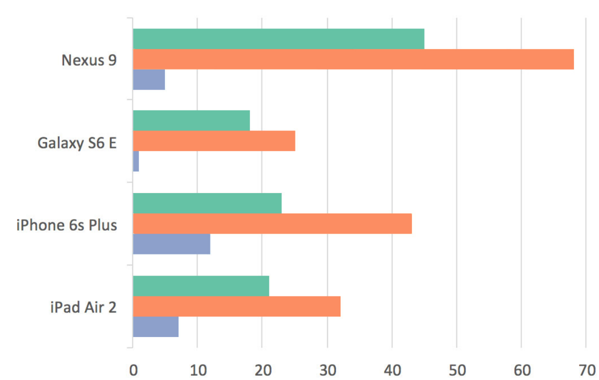


Figure 2: CPU usage in % in Unity (green), D3.js (orange) compared to idle mode (blue) [lower is better].

During the performance analysis it was very interesting to see, that the Nexus 9 tablet got noticeable warmer than the other devices. This effect mirrors in the device's CPU usage. However, no temperature measurements were carried out to investigate this effect. In general, less CPU usage is a big benefit from the perspective of smart devices because less energy consumption results in more battery time.

4.2 FPS analysis

The evaluation of the FPS data shows that Unity reaches a median of 57 FPS and D3.js version achieves a median of 51 FPS. Unity can be seen as the winner of this criteria of the performance comparison. The detailed median values of the evaluation part are shown in Figure 3.

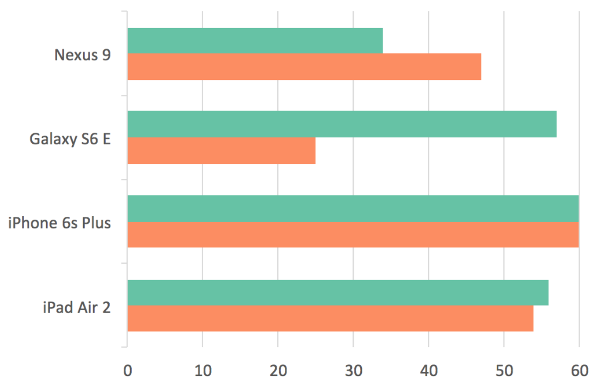


Figure 3: FPS rate while performing with Unity (green) and D3.js (orange) [higher is better].

It is very prominent, that the FPS rate of the D3.js version was pretty low on the Galaxy S6 Edge, despite the fact that the CPU usage on this device also stayed slightly. In contrast, the Nexus 9 tablet was the only device which reaches higher FPS with D3.js.

4.3 Loading Time Analysis

The result of the CSV data loading time measurement shows, that D3.js takes a median of 5.17ms. In contrast, Unity requires significantly more time for the raw data loading which results in a median of 15.17ms. Figure 4 shows the gap between both versions.

The measured time depends on the internal implementation of the loading methods of the visualizations which is the reason of the serious differences at the cycle times of these functions.

5 Conclusion

This study compared two different approaches for implementing cross-platform visualizations: cross-

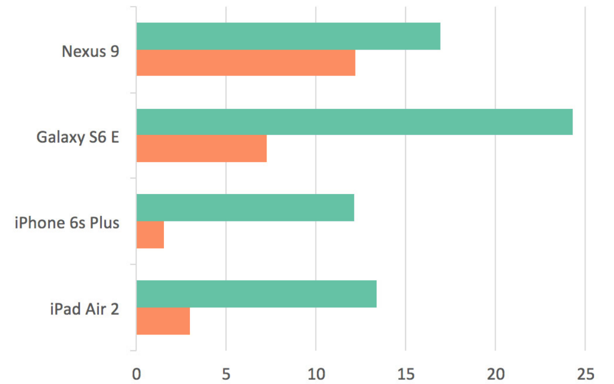


Figure 4: CSV loading times in ms while performing in Unity (green) and D3.js (orange) [lower is better].

compilation to native code and web technology, i.e. usage within a web browser.

For this, the Baby Name Explorer, as example of a realistic casual visualization design, was implemented in both Unity and D3.js. Our experimental comparison on four devices showed that FPS were comparable, D3.js was faster in initial data transformations, and Unity resulted in a lower CPU utilization.

In terms of developer experience, Unity's IDE supports C# as well as JavaScript for development. The cross-compilation and deployment of the Baby Name Explorer for all tested platforms worked seamlessly.

D3.js code is typically developed for a web environment. Due to the variety of web browsers, web based visualizations need to be tested on a wide selection before being released. During our experiment both implementations worked well.

Depending on our proof-of-concept, we demonstrated the benefits of the use of Unity for information visualization and cross-platform compilation in our field of research. In the next steps we will focus on the synchronization for collaboration and semantic zoom [WBR+16] and to show the ability to use this framework for visualization for the masses as called by Blumenstein et al. [BWA+15b] as an easy to use system.

Acknowledgements

This work was supported by the Austrian Science Fund (FWF) via the KAVA-Time and VisOnFire projects (no. P25489 and P27975), the Austrian Ministry for Transport, Innovation and Technology (BMVIT) under the ICT of the future program via the VALiD project (no. 845598) and under the Austrian Security Research Programme KIRAS via the project Courageous Community (no. 850196) as well as the project seekoi (no. 1154) funded by the Internet Foundation Austria (IPA).

References

[BLC12] Dominikus Baur, Bongshin Lee, and Sheelagh Carpendale. TouchWave: kinetic multi-touch manipulation for hierarchical stacked graphs. In *Proc. 2012 ACM int. conf. Interactive Tabletops and Surfaces, ITS*, pages 255–264. ACM, 2012.

[BNW+16] Kerstin Blumenstein, Christina Niederer, Markus Wagner, Grischa Schmiedl, Alexander Rind, and Wolfgang Aigner. Evaluating information visualization on mobile devices: Gaps and challenges in the empirical evaluation design space. In *Proc. 6th Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization, BELIV*, pages 125–132. ACM, 2016.

[BOH11] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. D3: Data-Driven Documents. *IEEE Trans. Vis. and Comp. Graphics*, 17(12):2301–2309, December 2011.

[BSB13] Enrico Bertini, Moritz Stefaner, and Dominikus Baur. Visualization on Mobile & Touch Devices. datastori.es podcast, `http://datastori.es/data-stories-25-mobile-touch-vis/`, 00:41:49 to 00:46:08, July 2013.

[BWA15a] Kerstin Blumenstein, Markus Wagner, and Wolfgang Aigner. Cross-Platform InfoVis Frameworks for Multiple Users, Screens and Devices: Requirements and Challenges. In *Workshop on Data Exploration for Interactive Surfaces DEXIS 2015*, pages 7–11, 2015.

[BWA+15b] Kerstin Blumenstein, Markus Wagner, Wolfgang Aigner, Rosa von Suess, Harald Prochaska, Julia Püringer, Matthias Zeppelzauer, and Michael Sedlmair. Interactive Data Visualization for Second Screen Applications: State of the Art and Technical Challenges. In *Proc. of the Int. Summer School on Visual Computing*, pages 35–48. Frauenhoferverlag, 2015.

[CM84] William S. Cleveland and Robert McGill. Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*, 79(387):531–554, 1984.

[CMS99] Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman. *Readings in Information Visualisation. Using Vision to Think*. Morgan Kaufmann, 1999.

[HBO10] Jeffrey Heer, Michael Bostock, and Vadim Ogievetsky. A tour through the visualization zoo. *Comm. of the ACM*, 53(6):59, 2010.

[HTA+15] Dandan Huang, Melanie Tory, Bon Adriel Aseniero, Lyn Bartram, Scott Bateman, Sheelagh Carpendale, Anthony Tang, and Robert Woodbury. Personal visualization and personal visual analytics. *IEEE Trans. Vis. and Comp. Graphics*, 21(3):420–433, March 2015.

[JJK08] Donald W. Johnson and T. J. Jankun-Kelly. A scalability study of web-native information visualization. In *Proc. Graphics Interface, GI*, pages 163–168, Toronto, 2008. Canadian Information Processing Society.

[KSC12] Daniel E. Kee, Liz Salowitz, and Remco Chang. Comparing interactive web-based visualization rendering techniques. In *Poster Proc. IEEE Conf. Information Visualization, InfoVis*, 2012.

[LAB+08] Tim Lammarsch, Wolfgang Aigner, Alessio Bertone, Silvia Miksch, Thomas Turic, and Johannes Gärtner. A comparison of programming platforms for interactive visualization in web browser based applications. In *Proc. 12th Int. Conf. Information Visualisation, iV*, pages 194–199, July 2008.

[LAMR14] Tim Lammarsch, Wolfgang Aigner, Silvia Miksch, and Alexander Rind. Showing important facts to a critical audience by means beyond desktop computing. In Yvonne Jansen, Petra Isenberg, Jason Dykes, Sheelagh Carpendale, and Dan Keefe, editors, *Death of the Desktop—Workshop co-located with IEEE VIS 2014*, 2014.

[LH14] Zhicheng Liu and Jeffrey Heer. The effects of interactive latency on exploratory visual analysis. *IEEE Trans. Vis. and Comp. Graphics*, 20(12):2122–2131, December 2014.

[MFB+07] J. D. Meier, Carlos Farre, Prashant Bansode, Scott Barber, and Dennis Rea, editors. *Performance testing guidance for web applications: patterns & practices*. Microsoft, United States?, 2007. OCLC: ocn245241921.

[Mun14] Tamara Munzner. *Visualization Analysis and Design*. A K Peters Ltd, 2014.

[PSM07] Zachary Pousman, John T. Stasko, and Michael Mateas. Casual Information Visualization: Depictions of Data in Everyday Life. *IEEE Trans. Vis. and Comp. Graphics*, 13(6):1145–1152, 2007.

[uni16] Unity – Game Engine, 2016. https://unity3d.com/.

[Wat05] Martin Wattenberg. Baby names, visualization, and social data analysis. In *Proc. IEEE Symp. Information Visualization, INFOVIS*, pages 1–7, October 2005.

[WBR+16] Markus Wagner, Kerstin Blumenstein, Alexander Rind, Markus Seidl, Grischa Schmiedl, Tim Lammarsch, and Wolfgang Aigner. Native cross-platform visualization: A proof of concept based on the Unity3D game engine. In *Proc. Int. Conf. Information Visualisation, iV*, pages 39–44. IEEE Computer Society Press, 2016.

Session 3: Information Visualization

Interactive Infographics in German Online Newspapers

Sandra Zwinger
University of Applied Sciences Burgenland
Eisenstadt, Austria
sandra.zwinger@live.at

Michael Zeiller
University of Applied Sciences Burgenland
Eisenstadt, Austria
michael.zeiller@fh-burgenland.at

Abstract

Interactive infographics are increasingly used in online journalism. Especially in data-driven journalism they provide a powerful tool to represent and communicate complex information. On the basis of a content analysis of five German-speaking online newspapers we show which kinds of interactive visualizations are preferred in German-speaking online media. However, users are still unfamiliar with interactive infographics. The results of a user survey demonstrate how readers of online newspapers use interactive infographics. These results point out how users interact with interactive infographics and how they assess the availability and findability of interactive infographics in online newspapers.

1. Introduction

Data-driven journalism (short: DDJ) collects, evaluates, interprets and presents large amounts of data. In an innovative and unique way data-driven journalism explains new insights and clarifies facts while telling complex stories on the basis of large amounts of retrieved data [Mat16] [GBC12]. Lorenz [Lor10] defines data-driven journalism as a workflow, where data is the basis for analysis, visualization and – most importantly – storytelling. The growth potential of data-driven journalism is vast and according to Weinacht und Spiller [WS14] it is one of the big issues in specialist publications on journalism and in education of journalists in Germany.

Data-driven journalism can help a journalist to tell a complex story through engaging information graphics [GBC12] [WR13]. In the reporting phase, visualizations can help journalists to identify themes and questions, to identify outliers or to find typical examples. In publishing, visualizations play multiple roles, e.g., to illustrate a point made in a story in a more compelling way, to remove unnecessary technical information from prose and to provide transparency about the reporting process to readers (especially interactive visualizations that allow exploration) [GBC12].

Information visualization is the use of interactive visual representations of abstract data to amplify cognition [Che10] [War12]. Information visualization in online media like online newspapers or online magazines can be interactive, i.e., it provides users with at least one option to control which and how much information shall be shown. Shneiderman [Shn96] proposed key principles for the success of information visualization that are based on his *Visual Information-Seeking Mantra*: Overview first, zoom and filter, then details-on-demand. Since then several types of interactivity have emerged. Information visualization provides several types of information graphics, e.g., functional infographics, cartographical graphics or visualization of statistical data, that can be enhanced with interactive elements.

Journalists tell stories based on their investigations. Data visualization is an appropriate communication medium for storytelling, in particular when the story also contains a lot of data [KM13]. However, these data stories told by data journalists differ from traditional forms of storytelling [SH10]. An emerging number of stories is enhanced with narratives including complex graphics and especially interactive graphics. Like the pioneers of data-driven journalism using interactive visualizations from Anglo-Saxon countries (e.g., *The Guardian*, *The New York Times*) the increasing number of data-driven journalistic projects in German-speaking countries use various visualizations with different options of interaction.

In this paper, we investigate which kinds of interactive infographics are used in data-driven journalism in Germany, Austria and Switzerland to effectively present complex data and causal relationships. A content analysis is applied to five examples of German-speaking online newspapers. The interactive infographics that have been published in 2014 and 2015 are analyzed.

Although interactive infographics are increasingly used in online media, readers really have to view them and use the control tools. First, readers have to identify

In: W. Aigner, G. Schmiedl, K. Blumenstein, M. Zeppelzauer (eds.): Proceedings of the 9th Forum Media Technology 2016, St. Pölten, Austria, 24-11-2016, published at http://ceur-ws.org

interactive infographics as such and the control tools have to be recognized intuitively. Thus, we investigate how readers of online newspapers assess the availability and findability of interactive infographics and which types of interaction they utilize.

2 Interactive Infographics

Information graphics (short: infographics) combine graphics, image and text to communicate information, data or knowledge effectively using graphic visual representations [Ale13] [Yau11] [Yau13]. Infographics aim at providing the percipient with new insights and a quick overview on complex facts on subjects like politics, science, technology, and nature that are hard to understand just using text-based information. They shall communicate complex issues efficiently and draw the attention of percipients to them. However, there is a continuous debate over visual embellishment, i.e., adding unnecessary visual embellishment – chart junk – versus a minimalistic approach using plain and simple charts [BMG+10]. Like the data-ink ratio for noninteractive infographics proposed by Tufte [Tuf01] to reduce chart junk, interactive infographics have to apply interactivity purposeful.

2.1 Types of Infographics

Infographics can be categorized in three basic types (Figure 1): [JS99] [Sta13]

- Principle representation
- Cartographic infographics
- Statistics chart

Figure 1: Three types of infographics (Source: [JS99])

The principle representation – also denoted as functional graphics – covers *What* and *How* questions and describes complex causal relationships in real or abstract form. Principle representations are subdivided in fact graphics, structure graphics and process graphics (Figure 2). [JS99]

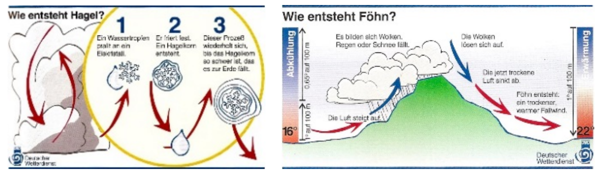


Figure 2: Fact graphics (left) und process graphics (right) (Source: [JS99])

The cartographic infographics (map) conveys space-oriented information in a clearly arranged, simple and understandable way that provides orientation. This type consists of event space map, topic map and weather map. [JS99]

Statistics charts help to illustrate quantities and compare them, especially large, complex sets of numbers and relations. This type consists of pie chart, bar chart, curve chart, area diagram, scatter plot, Sankey diagram and radar chart (Figure 3). [JS99] [Sta13]

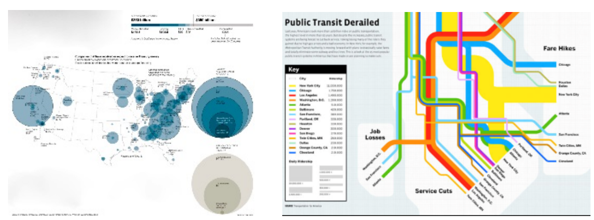


Figure 3: Bubble graph (scatter plot, left, Source: www.nytimes.com) and Sankey diagram (right, Source: www.ensbirasjon.blogspot.dk)

2.2 Interactions in Infographics

Many technical systems offer interactive performance. However, the interactive elements have to be recognized and utilized by users. This level of action is made up of a control dimension and a transmission dimension that allows the system to react on user input. The control dimension ist subdivided in options for selection (selection of existing content, e.g., click on a hyperlink) and options for modification (change of system range by input; e.g., input of text) [QS06].

There exist several methods of interaction to manipulate a visual representation, like scrolling, overview plus detail, or focus plus context, filtering or data reordering [Maz09].

Weber and Wenzel [WW13] define interactive infographics as being a visual representation of information that integrates several modes (at least two), e.g., image/video, spoken or written text, audio, layout, etc. (image mode is constitutive), to a coherent ensemble that offers at least one option of control to the user. The provided option of control can be, e.g.,

- Start or Stop button
- Forward or backward button
- Menu item to select
- Timeline or time controller
- Filter, data request or input box

2.3 Classification of Interactive Infographics

While evaluating interactive infographics we analyze five features which cover interaction as well as narrative issues: degreee of interactivity, course of action, communicative intent, "W-questions" and topic. We follow a typology introduced by Weber and Wenzel [WW13] and Weber [Web13]. Other features like *genre* or *visual narrative* as in [SH10] might be applied as well, but have not been included.

2.3.1 Degree of Interactivity

Weber and Wenzel [WW13] distinguish three degrees of interactivity:

- Low interactivity
- Medium interactivity
- High interactivity

A low level of interactivity allows users to navigate within the infographics and select content, e.g., by using internal links, zooming, mouseover effects for showing details, Next or Start buttons, but without changing the infographics. On a medium level of interactivity users can manipulate the infographics, e.g., by a timeline slider or menu items, thus showing changes and comparing information. A high level of interactivity enables users to explore the infographics and to interact with data and information, e.g., by inputs, filtering or data retrieval (Figure 4).

We analyze only the degree of interactivity, but do not identify the ways of manipulation as in [SH10], like hover highlighting and details, filtering and selection, search or navigation buttons.

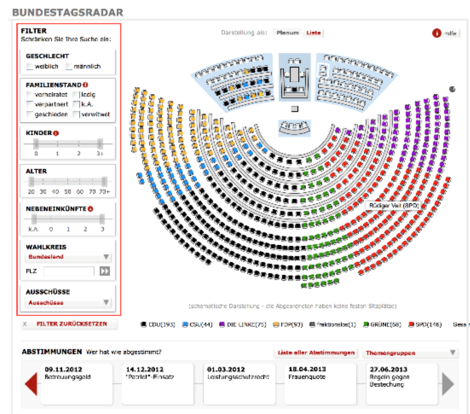


Figure 4: Infographics showing German House of Parliament with high interactivity (Source: www.spiegel.de, [WW13])

2.3.2 Course of Action

Weber and Wenzel [Web13] [WW13] distinguish three types of course of action:

- Linear
- Nonlinear
- Linear-nonlinear

Linear interactivity enables the user to move (forward or backward) through a predetermined linear sequence [Sim97]. The linear type is based on a step-by-step course defined by the author (i.e., author-driven [SH10]). The user follows a strict path and does not have to explore the visualization by himself. Navigation tools like *Start*, *Stop*, *Forward*, *Backward* or *Next* are used to navigate in a liner course. A navigation bar provides a better overview on the whole sequence. [Web13]

A nonlinear visualization does not provide a prescribed ordering and requires a high degree of interactivity by the user [Rya06] – its narrative being reader-driven [SH10]. Nonlinear-based interactive infographics provide the user with many ways to explore and query the visualization – including free exploration without predefined navigation paths. Navigation tools for nonlinear infographics include filter, input box, data query or brushing (various visual representations of same data). [Web13]

The linear-nonlinear type results from a hybridization of the linear and nonlinear course of action. This type is a hybrid of the author-driven and reader-driven approach [SH10] that enables the author to communicate his message using a predefined path, but still allowing the user a certain amount of selection ability. Navigation tools for linear-nonlinear infographics include interactive timelines, time controller, and integrated navigation menu. [Web13]

2.3.3 Communicative Intent

Based on Nichani und Rajamanickam's taxonomy [NR03] we use the following categories to represent the communicative intent oft interactive infographics:

- Narratives: Stories told from a distinct point of view (e.g., anecdotes, personal stories, business stories, case studies).
- Instructives: Step-by-step instructions explaining how things work or how events occur.
- Exploratives: Allow readers to discover the intent themselves by active exploring and sensemaking.
- Simulatives: Allow readers to experience the intent (usually a real world phenomena) themselves.

2.3.4 W-Questions

In addition to the communicative intent [NR03] the communicative function can be derived using an extended version of the journalistic W-questions [Web13] [Roa09]:

- What/Who
- When
- Where
- How
- Why
- How much

2.3.5 Topic

Following the news departments of a newspaper/ magazine that classify a newspaper article we classify the topic of the interactive infographics:

- Politics/Economics
- Accidents/Natural disaster
- Consumption
- Sports
- Science/Society
- Crime
- Others

3 Analysis of Utilization of Interactive Infographics in German-speaking Online Newspapers

Interactive infographics have various occurrences and are used for multiple purposes. However, not all types of interactive infographics can be utilized in data-driven journalism. In this section we analyze which interactive infographics are used in data-driven journalism in Germany, Austria and Switzerland to efficiently present complex data and their causal relationships. [Zwi16a]

3.1 Method

The analysis of interactive infographics of selected online newspapers is based on evaluation research. The object of the evaluation – the interactive infographics – is assessed according to predefined goals and specified criteria. [GJ14] [SM14]

To represent the German-speaking online newspapers we selected five newspapers from Germany, Austria and Switzerland that are published in print and online: „Zeit Online“, „Berliner Morgenpost“, „Neue Zürcher Zeitung“, „Der Standard“ und „Kurier”. The objects to be evaluated are the interactive infographics that had been published between January 1, 2014 and December 31, 2015 in the newspapers' online portals.

A total of 276 interactive infographics have been identified and subsequently analyzed: Zeit Online 36, Berliner Morgenpost 54, Neue Zürcher Zeitung 89, Der Standard 60, and Kurier 37 infographics.

Based on the typology and classification introduced in section 2 a list of criteria is defined. The type of infographics is classified on a nominal scale (scale values A-M and A-E respectively) (Table 1). The classification criteria are assessed using an ordinal scale with varying degree (Table 2).

3.2 Results

3.2.1 Example: *Zeit Online*

As an example, we present the results of the evaluation of interactive infographics that have been published by „Zeit Online” [Zwi16a]. The other newspapers, i.e., „Berliner Morgenpost“, „Neue Zürcher Zeitung“, „Der Standard“ und „Kurier”, have been evaluated in the same way.

Table 1: Scale values for type

Criterion	Value
Type of infographics	Definitely a single type of infographics: 0: no, 1: yes
	If type is 1⇒ Single type: A: Fact graphics (view, magnifier, section) B: Structure graphics (building plan, text image) C: Process graphics (overall view, sequence) D: Event space map E: Topic map F: Weather map G: Pie chart H: Bar chart I: Curve chart, area diagram J: Scatter plot, bubble graph K: Sankey diagram L: Radar chart M: Other
	If type is 0 ⇒ Hybrid type: A: Principle representation with cartographic IG B: Principle representation with statistics chart C: Cartographic infographics with statistics chart D: Mixed statistics chart E: Other

Table 2: Criteria for classification

Criterion	Value
Degree of interactivity	0: Low interactivity (internal links, zooming, mouseover, Next button) 1: Medium interactivity (menu items, timeline slider) 2: High interactivity (filter, input box, brushing)
Course of action	0: Linear (step-by-step; Forward, Backward or Next) 1: Nonlinear (filter, input box, data query, brushing) 2: Linear-nonlinear (interactive timeline, time controller, integrated navigation menu)
Communi-cative intent	0: Narratives 1: Instructives 2: Exploratives 3: Simulatives
W-Questions	0: What/Who 1: When 2: Where 3: How 4: Why 5: How much
Topic	0: Politics/Economics 1: Accidents/Natural disaster 2: Consumption 3: Sports 4: Science/Society 5: Crime 6: Others

We identified 36 interactive infographics that have been published during the period of investigation by "Zeit Online". 27 of them are single type infographics, and 9 are hybrid infographics. The single type infographics are made up of seven topic maps, four bar charts, four curve charts, three event space maps, two fact graphics and one process graphics; the remaining six infographics are of type *Other*. The nine hybrid infographics consist of three cartographic infographics with statistics chart, two principle representation with statistics chart, two mixed statistics charts and two infographics of type *Other*.

"Zeit Online" published 18 interactive infographics with a low degree of interactivity, 13 infographics with medium interactivity and 5 infographics with high interactivity. In category *Course of action* 6 infographics were linear, 25 infographics were nonlinear and 5 of the combined type linear-nonlinear. Analyzing the communicative intent shows 21 infographics of type *Exploratives*, 10 infographics of type *Instructives*, 3 of type *Simulatives* and 2 of type *Narrratives*.

Most infographics have been published relating to topics in science and society: 19 infographics in total. 6 infographics cover politics and economics, 5 deal with sports, 5 deal with consumption and one infographics is on crime (Figure 5).

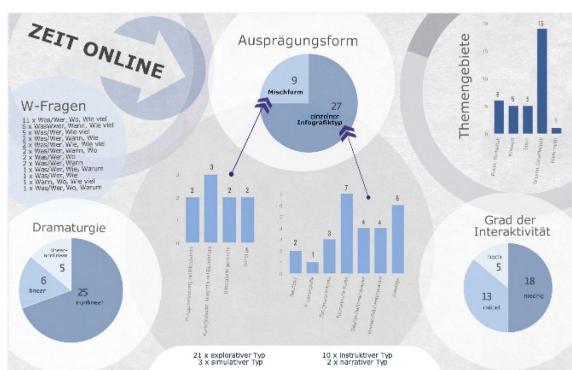


Figure 5: Results „Zeit Online“ (Source: [Zwi16a])

3.2.2 Using Interactive Infographics

The unrepresentative evaluation research analyzed 276 interactive infographics in five selected online newspapers (period: January 1, 2014 until December 31, 2015) showing that in German-speaking data-driven journalism typically interactive infographics are used that belong to single type infographics (77.54 % vs. 22.46 % of hybrid type). Most commonly used are bar charts (23.83 %) and curve charts/area diagrams (18.69 %). Topic maps are used at a percentage of 17.76 % and event space maps are used at a percentage of 12.15 %. Rarely used are structure graphics (2.34 %), pie charts (1.87 %), Sankey diagrams (1.87 %), process graphics (1.40 %) and fact graphics (1.40 %) (Figure 6).

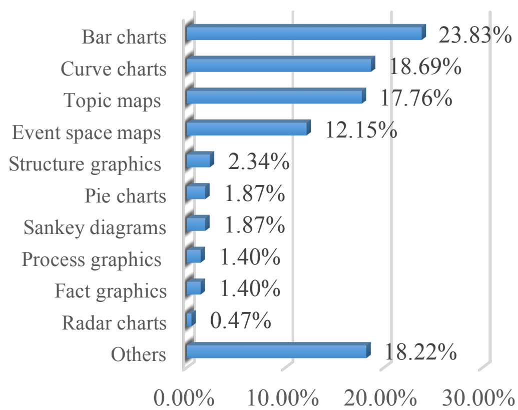


Figure 6: Use of single type infographics (Source: [Zwi16a])

The majority of interactive infographics, i.e., 43.11 %, shows a low degree of interactivity. 38.77 % of the analyzed infographics have a medium degree of interactivity, whereas only 18.12 % of the analyzed infographics offer a high degree of interactivity (Figure 7).

The vast majority of published interactive infographics exhibits a nonlinear course of action. 75.0 % are nonlinear infographics, whereas 13.77 % of the infographics are identified as belonging to the linear-nonlinear type. Only 11.23 % of the analyzed infographics have a linear course of action (Figure 8).

The evaluation of the communicative intent depicts that *Exploratives* are used most at a ratio of 70.65 %. We identified 21.02 % *Instructives*. *Narratives* (5.07 %) and *Simulatives* are seldom used (Figure 9).

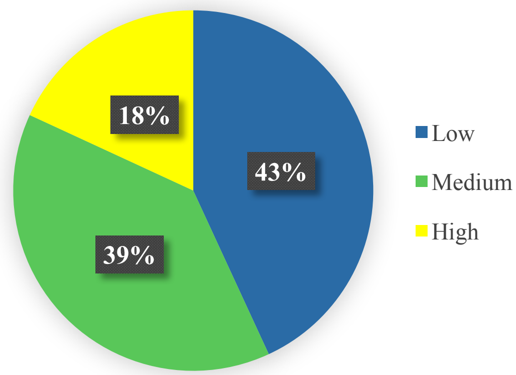


Figure 7: Degree of interactivity (Source: [Zwi16a])

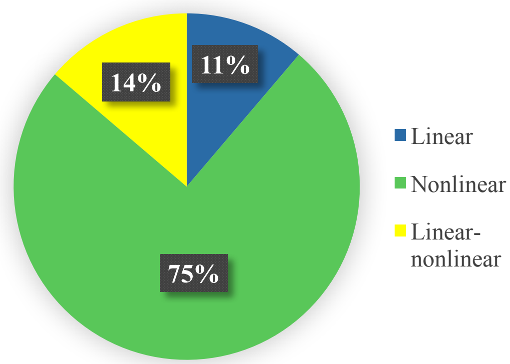


Figure 8: Course of action (Source: [Zwi16a])

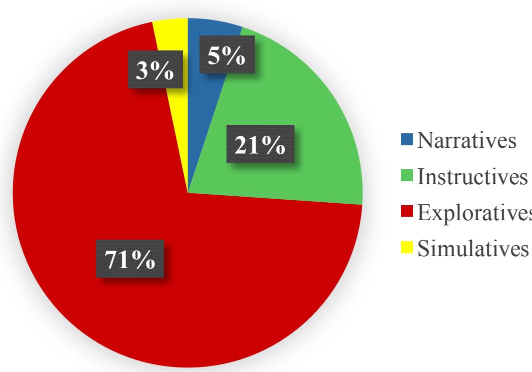


Figure 9: Communicative intent (Source: [Zwi16a])

Interactive infographics are most frequently used in online newspapers to illustrate information on the topic science and society with 46.01 % of all occurrences. 34.06 % of the analyzed infographics cover issues on politics and economics. 10.51 % cover sports, 6.16 % crime and 2.54 % consumption. Accidents and natural disaster are seldom discussed with the help of infographics (0.36 %). Another 0.36% could not be assigned clearly and were labeled as *Other* (Figure 10).

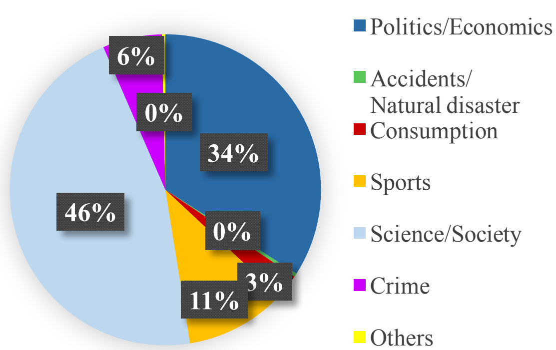


Figure 10: Topics (Source: [Zwi16a])

4 Utilization of Interactive Infographics by Readers

During the evaluation research process it turned out that interactive infographics are hard to identify in online newspapers. In particular, it was difficult to find infographics in the two Austrian online newspapers. Most newspapers have no dedicated sections aggregating interactive infographics and even the integrated search function often failed.

In this chapter we analyze how Austrian readers of online newspapers assess the availability and findability of interactive infographics and how they use the facilities for interaction. Our goal is the demonstration that readers of online newspapers search purposeful for interactive infographics or just use them by pure chance. We examine how laborious and time-consuming readers estimate searching for infographics. The intensity of the use of infographics and the frequency of the utilization of the individual facilities for interaction are determined.

4.1 Method

To determine the opinion of readers of online newspapers a quantitative research approach has been chosen [BB14]. Since the target group is very familiar with online tools we chose to conduct a nonrepresentative online survey [WH14]. The online survey was made up of 24 questions that included closed questions (single choice, multiple choice questions) as well as matrix questions with different evaluation scales. [MG14]

Only readers of online newspapers have been surveyed. Persons that got access to the survey, but do not read online newspapers were identified at the beginning using a knockout question. The survey was published in the off-topic forum of the Austrian newspaper “DerStandard”, on the Facebook page of the Austrian newspaper “Kurier”, and additionally on the Facebook account of one of the authors and among students of the University of Applied Sciences Burgenland [Zwi16b]. Survey period: June and July 2016.

259 persons participated in the survey. Due to the research design primarily Austrian readers of German-speaking online newspapers have been surveyed. 215 persons (83.01 %) answered the preceding knockout question positive and were identified as readers of online newspapers. The following outcome of the survey refers to this group.

4.2 Results

86.98 % of the respondents (187 persons of 215) deliberately view interactive infographics. Asked for the reasons of using those infographics they name the well-known advantages of infographics like illustration of all relevant numbers/facts (77.54 %), a clearly structured overview (63.64 %) or good memorability due to the combination of text and image (51.34 %). Table 3 shows the results of this question (multiple references possible).

The respondents could rate on a five-point Likert scale (“very intensive” to “less intensive”) how intensively they use the offered facilities to interact with the infographics. More than half of the persons questioned use them “moderately intensive” (54.01 %, 101 persons). The infographics are “very intensively” used by only 6.42 % respondents (12 persons). 25.13 % (47 persons) used them “intensively”. The option “rather less intensive” has been named by 13 persons (6.95 %) and there are 14 users who use the infographics “less intensive” (Fig. 11).

Table 3: Reasons for using interactive infographics (Source: [Zwi16b])

Answer option	Percentage	#
Illustration of all relevant numbers/ facts	77.54 %	145
Clearly structured overview put into graphs	63.64 %	119
Good memorability due to the combination of text and image	51.34 %	96
Relevant information can be filtered and visualized	34.22 %	64
Easy search for data/information	26.74 %	50
Do not like to read long text	21.93 %	41
Offers to view data/information in various representations	16.04 %	30
Other	1.60%	3

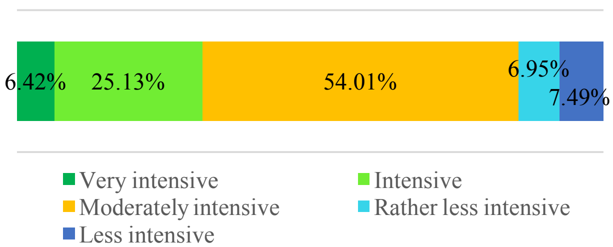


Figure 11: Intensity of using the facilities of interaction (Source: [Zwi16b])

Interactive infographics offer different functions to control the graphics. Table 4 presents how frequently (or seldom) users apply those control tools.

Table 4: Frequency of use of control tools

n=187	Very frequently	Frequently	Occasionally	Seldom	Never
Start, Stop button	11,23 % 21	25,67 % 48	35,83 % 67	22,46 % 42	4,81 % 9
Forward, Backword button	10,70 % 20	29,95 % 56	39,57 % 74	18,18 % 34	1,60 % 3
Navigation bar (menu)	26,20 % 49	46,52 % 87	19,25 % 36	6,42 % 12	1,60 % 3
Filter	20,86 % 39	36,36 % 68	26,74 % 50	13,37 % 25	2,67 % 5
Timeline slider	9,09 % 17	25,13 % 47	37,97 % 71	21,93 % 41	5,88 % 11
Input box	3,21 % 6	18,18 % 34	35,29 % 66	32,09 % 60	11,23% 21

The respondents have been asked whether those control functions are sufficiently recognizable and easy to identify. Only 2.67 % of the respondents (5 persons) consider the control tools as “very well recognizable”. 33.16 % (62 persons) consider them as “good recognizable” and the majority of 37.97 % (71 respondents) consider the control tools as “moderately recognizable”. 20.32 % respondents (38 persons) have not been satisfied with the visibility of the control tools and graded them as “poorly recognizable”. One respondent (0.53 %) did not find the control tools and answered “not recognizable” (10 persons, i.e., 5.35 %, did not specify) (Fig. 12). Obviously there is need for action.

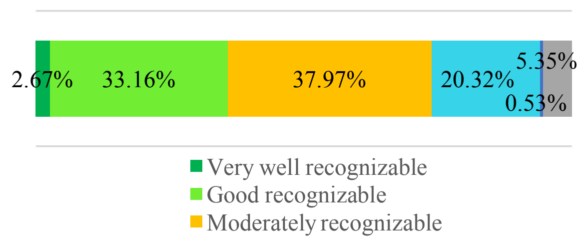


Figure 12: Perceptibility of control tools (Source: [Zwi16b])

Only 24.06 % of the respondents (45 persons) indicated that they actively search for interactive infographics in online newspapers. These persons would see a big advantage if online newspapers mark or label interactive infographics or offer easy access using dedicated menu items. 75.94 % (142 respondents) do not search actively for articles that contain interactive infographics. However, a majority of three-quarters of those respondents (n=142; i.e., 106 persons) mentioned that they would use infographics more often, if they would be easier to find.

53.48 % respondents of the survey quoted that they think that currently most interactive infographics deal with politics. 24.60 % respondents indicated that most interactive infographics are about economics. The respondents would prefer if more interactive infographics are published concerning science (24.60 %), economics (22.99 %), consumption (10.70 %), crime (9.63 %) and politics (8.56 %) (remark: only single response allowed). However, 73.26 % of the respondents of the survey would appreciate an increasing offering of interactive infographics in online newspapers.

5 Conclusion

The evaluation of interactive infographics that have been published in the years 2014 and 2015 in five selected German-speaking online newspapers shows inhomogeneous and diverse results. The analyzed infographics depict that data-driven journalism in German, Austria and Switzerland typically makes use of infographics that belong to a single type (77.54 %). Most commonly used are bar charts at a percentage of 23.83 %. Often single type infographics are combined to hybrid forms that are used at a percentage of 22.46 %. Most examples of those hybrid forms (53.23 %) are cartographic infographics with statistics chart.

The majority of interactive visualizations only have a low degree of interactivity (43.11 %). Users can manipulate the infographics easily, e.g., by using *Next* or *Start* buttons or mouseover effects for showing additional information. Only a small percentage of 18.12 % of the analyzed infographics offer a high degree of interactivity and allow users to interact with the help of input boxes, data retrieval or filtering.

Three-quarters of the analyzed infographics have a nonlinear course of action. They offer several options to explore and manipulate the infographics. By means of navigation tools like input box, filter, menu item or mouseover effects the requested information can be displayed.

The survey of readers of German-speaking online newspapers demonstrates that the offering of interactive infographics is highly accepted. However, it is often hard to recognize that a graphics is an interactive infographics with corresponding control tools. Consequently, those control tools are only moderately or seldom used. There is potential for improvement to identify controls for interactivity more clearly and thus allowing users to fully utilize the offer of information of the infographics. For example the findability of interactive infographics might be improved by dedicated sections in the online newspapers aggregating infographics (e.g., specific menu items in the main navigation) or by others ways of identification. The reason why interactive infographics are often not used is due to the simple fact that they cannot be found. However, readers of online newspapers have strong interest in using and exploring interactive infographics.

6 Summary

Data-driven journalism deals with the collection, evaluation, interpretation and presentation of large amounts of data. Frequently, visualizations are used to present these data in online reporting. Online media allow to design those infographics interactively. Users can determine by themselves which and how many information shall be shown. For example, interactive infographics with nonlinear course of action provide a high degree of interactivity up to free exploration without predefined navigation paths.

Based on the evaluation of 276 interactive infographics that have been published in 2014 and 2015 in five German-speaking online newspapers we demonstrated which types of visualization are used in data-driven journalism in Germany, Austria and Switzerland. The evaluation research shows a large variety of used types of infographics. At the same time, we recognize a great potential for improvement when applying those infographics in data-driven journalism. In particular, the degree of interactivity can be improved to facilitate even more flexibility when exploring those infographics.

From the point of view of readers of online newspapers we identified potential for improvement concerning the findability or recognizability of interactive infographics. Additionally, various control tools for the manipulation of infographics are available, but they are used only to a limited extend due to the fact that they cannot be identified conveniently. Therefore, improvements in the usability of interactive infographics might increase the utilization of infographics. Usability issues have not been investigated in the present analysis, but further research should examine the usability of interactive infographics in data-driven journalism.

References

[Ale13] K. Alexander. *Kompedium der visuellen Information und Kommunikation*. Springer Vieweg, 2013.

[BB14] N. Baur, J. Blasius. Methoden der empirischen Sozialforschung: Ein Überblick. In N. Baur, J. Blasius (Hrsg.). *Handbuch Methoden der empirischen Sozialforschung*. Springer VS, 41-62, 2014.

[BMG+10] S. Bateman, R. Mandryk, C. Gutwin, A. Genest, D. McDine, C. Brooks. Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts. In *ACM*

Conference on Human Factors in Computing Systems (*CHI 2010*), 2573-2582, 2010. DOI=10.1145/1753326.1753716.

[Che10] C. Chen. *Information Visualization: Beyond the Horizon*. 2nd Ed., Springer, 2010.

[GBC12] J. Gray, L. Bounegru, L. Chambers (Eds.). *Data Journalism Handbook*. European Journalism Centre and the Open Knowledge Foundation. http://datajournalismhandbook.org/

[GJ14] M. Gollwitzer, R. Jäger. *Evaluation kompakt*. Beltz, 2nd Ed., 2014.

[JS99] A. Jansen, W. Scharfe. *Handbuch der Infografik: Visuelle Information in Publizistik, Werbung und Öffentlichkeitsarbeit*. Springer, 1999.

[KM13] R. Kosara, J. Mackinlay. Storytelling: The Next Step for Visualization. *IEEE Computer* (Special Issue on Cutting-Edge Research in Visualization), 46(5), 44–50, 2013.

[Lor10] M. Lorenz. Status and outlook for data-driven journalism. In *Data-driven journalism: what is there to learn?* European Journalism Centre, 8-17, 2010.

[Mat16] L. Matzat. *Datenjournalismus: Methode einer digitalen Welt*. UVK Verlagsgesellschaft, 2016.

[Maz09] R. Mazza. *Introduction to Information Visualization*. Springer, 104-123, 2009.

[MG14] H.D. Mummendey, I. Grau. *Die Fragebogen-Methode*. Hogrefe, 6th Ed., 2014.

[NR03] M. Nichani, V. Rajamanickam. Interactive Visual Explainers – A Simple Classification. *elearningpost*, September 2003. http://www.elearningpost.com/articles/archives/interactive_visual_explainers_a_simple_classification/

[QS06] O. Quiring, W. Schweiger. Interaktivität – ten years after. *Medien & Kommunikationswissenschaften*, 54 (1):1-20, 2006.

[Roa09] D. Roam. *The Back of the Napkin: Solving Problems and Selling Ideas with Pictures*. Portfolio.

[Rya06] ML. Ryan. *Avatars of Story*. University of Minnesota Press.

[SH10] E. Segel, J. Heer. Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139-1148, November 2010.

[Shn16] B. Shneiderman. The Eyes Have It: a task by data type taxonomy for information visualization. Proceedings of the 1996 *IEEE Symposium on Visual Languages*, 336-343, 1996.

[Sim97] R. Sims. Interactivity: A Forgotten Art? *Computers in Human Behavior*, 13(2):157-180, 1997.

[SM14] R. Stockmann, W. Meyer. *Evaluation. Eine Einführung*. 2nd Ed., UTB, 2014.

[Sta13] T. Stapelkamp. *Informationsvisualisierung: Web-Print-Signaletik*. Springer Vieweg, 2013.

[Tuf01] E. Tufte. *Visual Display of Quantitative Information*. 2nd Ed., Graphics Press, 2001.

[War12] C. Ware. *Information Visualization: Perception for Design*. 3rd Ed., Elsevier, 2012.

[WBT13] W. Weber, M. Burmester, R. Tille. *Interaktive Infografiken*. Springer Vieweg, 2013.

[Web13] W. Weber. Typen, Muster und hybride Formen: Ein Typologisierungsmodell für interaktive Infografiken.. In W. Weber, M. Burmester, R. Tille. *Interaktive Infografiken*. Springer Vieweg, 25-37, 2013.

[WH14] P. Wagner, L. Hering. Online-Befragung. In N. Baur, J. Blasius (Hrsg.). *Handbuch Methoden der empirischen Sozialforschung*. Springer VS, 661-673, 2014.

[WR13] W. Weber, H. Rall. "We Are Journalists." Production Practices, Attitudes and a Case Study of the New York Times Newsroom. In W. Weber, M. Burmester, R. Tille. *Interaktive Infografiken*. Springer Vieweg, 161-172, 2013.

[WS14] S. Weinacht, R. Spiller. Datenjournalismus in Deutschland: Eine explorative Untersuchung zu Rollenbildern von Datenjournalisten. *Publizistik*, 59(4):411–433, December 2014. DOI 10.1007/s11616-014-0213-5.

[WW13] W. Weber, A. Wenzel. Interaktive Infografiken: Standortbestimmung und Definition. In W. Weber, M. Burmester, R. Tille. *Interaktive Infografiken*. Springer Vieweg, 3-23, 2013.

[Yau11] N. Yau. *Visualize This: The FlowingData Guide to Design, Visualization, and Statistics*. Wiley Publishing, 2011.

[Yau13] N. Yau. *Data Points: Visualization That Means Something*. John Wiley & Sons, 2013.

[Zwi16a] S. Zwinger. *Visualisierung im deutschsprachigen online Datenjournalismus*. Bachelor thesis, University of Applied Sciences Burgenland, February 2016.

[Zwi16b] S. Zwinger. *Interaktive Infografiken in österreichischen Onlinezeitungen*. Bachelor thesis, University of Applied Sciences Burgenland, August 2016.

Exploring Media Transparency With Multiple Views

Alexander Rind,[1,2] David Pfahler,[1] Christina Niederer,[2] and Wolfgang Aigner[1,2]

[1] TU Wien, Austria
[2] St. Poelten University of Applied Sciences, Austria

Abstract

Politically concerned citizens and data journalists want to investigate money flows from government to media, which are documented as open government data on 'media transparency'. This dataset can be characterized as a dynamic bipartite network with quantitative flows and a large number of vertices. Currently, there is no adequate visualization approach for data of this structure. We designed a visualization providing coordinated multiple views of aggregated attribute values as well as short tables of top sorted vertices that can be explored in detail by linked selection across multiple views. A derived attribute 'trend' allows selection of flows with increasing or decreasing volume. The design study concludes with directions for future work.

1 Introduction

Independent news and media are a cornerstone of modern democracy – often called the fourth power. However, governmental advertisement and sponsorships could influence news coverage limiting the media's independence. In Austria, the federal law on Transparency in Media Cooperation and Funding [Med15] makes it mandatory to disclose such flows of money from legal entities (e.g., federal ministries, cities, economic chambers, government-owned companies) to media institutions (e.g., newspaper, TV, radio, online). The Austrian Regulatory Authority for Broadcasting and Telecommunications [RTR] collects these data and makes them publicly available via the Austrian open government data portal [RTR16].

In: W. Aigner, G. Schmiedl, K. Blumenstein, M. Zeppelzauer (eds.): Proceedings of the 9th Forum Media Technology 2016, St. Pölten, Austria, 24-11-2016, published at http://ceur-ws.org

This so-called media transparency (MT) dataset is a valuable resource for politically concerned citizens as well as for data journalists [Aus15, Lor10]. They are interested in exploring the available data independently looking for stories beyond prearranged summary statistics. However, the MT dataset is much too large to be browsed line by line. Neither is it sufficient to look only at the largest flows of money because many possible questions of interest focus on changes over time and the many-to-many relationship between legal entities and media [NRA$^+$16]. For this purpose it is useful to conceptualize the MT dataset's money flows as time-dependent attributes on the edges of a bipartite network. Simple data analysis tools such as spreadsheets do not adequately support such a data structure.

Interactive visual representations of data [CMS99, Mun14] are a well-suited approach to explore complex datasets. Many visualization techniques have demonstrated their value in exploring time-oriented data [AMST11] and network data [BBDW16, HSS15]. However, the combination of time with quantitative flows in a bipartite networks is still an open challenge for visualization research [NAR15].

This paper contributes a visualization design study [SMM12] for time-oriented quantitative flows in a bipartite network. It uses the MT dataset as example and non-expert users such as citizens and journalists as target audience. After surveying related work in Section 2 and characterizing the domain problem in Section 3, we present the justified visualization design in Section 4. Next, a usage scenario demonstrates the design's utility in Section 5. The paper concludes with reflections for future development.

2 Related Work

The design space of network visualization has been mapped in some recent state-of-the-art reports: Hadlak et al. [HSS15] identified five facets of concern for visualizing a network: (i) its structure comprised of nodes and edges, (ii) partitions, (iii) the attributes of nodes and edges, (iv) dynamics, i.e., change over

Table 1: Raw format of the media transparency (MT) dataset (first four entries)

legal entity	time	law	medium	money amount
Abfallwirtschaft Tirol-Mitte GesmbH	Q4/2012	§2	Bezirksblätter Tirol	8,122.32 €
Agrarmarkt Austria Marketing GesmbH	Q4/2012	§2	Falstaff	26,418.00 €
Agrarmarkt Austria Marketing GesmbH	Q4/2012	§2	Connoisseur Circle	6,142.50 €
Agrarmarkt Austria Marketing GesmbH	Q4/2012	§2	bz-Wiener Bezirkszeitung	7,031.16 €
	⋮			

time, and (v) spatialization such as geographic context of nodes. Beck et al. [BBDW16] addressed in particular visual representations for dynamic networks such as animation and timeline. Von Landesberger et al. [vLKS+11] focused on large networks. Niederer et al. [NAR15] surveyed visualization of dynamic, weighted and directed networks, and thus, data of a structure similar to the MT dataset.

Examples of such related visualization designs are DOSA [vdEvW14], egoSlider [WPZ+16], egoLines [ZGC+16], Graph Comics [BKH+16], TimeArcTree [GBD09], and Visual Adjaceny List [HBW14]. However, none of these approaches explicitly considers the bipartite nature of the MT dataset, i.e., that there are distinct nodes for legal entities and for media.

We could identify only one scholarly work focusing on visualizing the MT dataset in particular: Niederer et al. [NRA+16] investigated the visualization needs of data journalists based on four interviews that were anchored on the MT dataset as exemplary scenario. Besides that, there is some press coverage on the data and some articles are accompanied by interactive web infographics (e.g., derStandard.at [Ham], Paroli Magazin [Lan]). Yet, these infographics present a subset of the available data that has been aggregated and filtered to support their articles' story. Since they allow only minimal interactivity, further exploration is not possible. Furthermore, since 2013 the open source software project Medientransparenz Austria [SBSV] provides an interactive online tool that shows the complete MT dataset. It integrates several visual representations giving insight into the data, but its views require much scrolling and are distributed across multiple pages. In addition, changes of money flow over time are not explicitly represented.

3 Background

As a fundament for developing a novel visualization design for the MT dataset, we must first understand its background and characterize the domain problem.

The law [Med15] regulates three categories of money flows that need to be disclosed: §2 covers advertisement, §4 sponsorships, and §31 ORF programme fees. Each quarter, each legal entity is obligated to make a disclosure for both §2 and §4. Every media cooperation involving more than 5,000 € needs to be included with the recipient's name and the amount of money accumulated in the quarter. If a legal entity had no such media cooperation, it still has to submit a nil report.

The MT dataset is published on an open data portal [RTR16] each quarter of a year with data covering the preceding eight quarters. The raw data are formatted as semicolon-separated values in a text file. Table 1 shows the five relevant variables: name of the legal entity, time specified by year and quarter, category of legal background, name of the medium, amount of money (quantitative). Additionally, the raw data contains a variable that flags nil reports.

3.1 Data Abstraction

The MT dataset is comprised of the quarterly money transferred from legal entities to media. We can conceptualize these data as time-dependent flows in a bipartite network (Figure 1) [NRA+16]. The network's underlying graph is bipartite because its vertices can be divided in two disjoint sets – legal entities and media – and each edge connects vertices of different sets. These edges are directed and weighted representing the flow of money from legal entities to media. The network is dynamic both in terms of its structure (vertices and edges can appear or disappear over time) and its quantitative flows (weights changing over time) [vLKS+11]. The time-oriented aspect of the data can be characterized as instants on a discrete, interval-

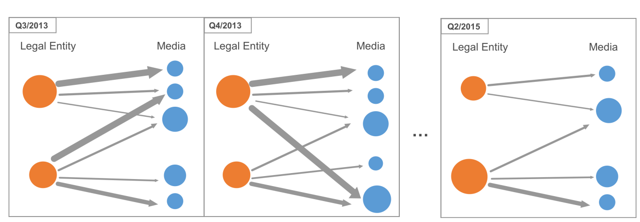


Figure 1: Conceptualizing the MT dataset as time-dependent flows in a bipartite network
Figure by [NRA+16] used with permission.

based, linear time domain with the granularities quarter and year [AMST11].

This abstract data structure has some benefits over the raw data's table structure: The central aspects of the problem domain (legal entities, media, and flows) are represented directly as data items, which can have properties from derived data such as aggregated money. Network metrics such as in-degree can be examined. They can also be manipulated by user interaction.

3.2 Preprocessing and Analysis of Data Scale

We perform some preprocessing to achieve a better data basis for our visualization:

(1) We substitute the original MT dataset with data from the Medientransparenz Austria project [SBSV], which have two benefits: First, they have included data for all quarters since the start of the MT dataset in Q3/2012. Second, they have preprocessed the data to clean different forms of writing the names of media and legal entities. Such inconsistencies could result either from typos or from the organization actually being renamed.

(2) Next, we discard nil reports from the data. Even though these nil reports make up about 80% of all records, they cannot add any insight to our design as they have missing values for media name and amount of money.

(3) Finally, we also discard programme fees (legal category §31) because on the one hand there are only one or two records per quarter and on the other hand their amount is much higher than any other record. The median §31 amount is about 80 times as much as the highest regular amount.

As of summer 2016, the preprocessed MT dataset encompasses 36,261 quarterly money flows over 15 quarters (Q3/2012–Q1/2016). So that one quarter has on average circa 2,400 flows. 34,717 flows (96%) have §2 as legal background and there are 1,544 flows for §4. (30 flows for §31 have been discarded.)

There are 993 distinct legal entities and 3,813 distinct media. Legal entities have between 1 and 1,782 outgoing flows (median = 8; average = 36.2), if we count each quarter as a separate flow. These flows connect them to between 1 and 618 distinct media (median = 3; average = 12.2). 71 legal entities maintain a continuous flow over all 15 quarters to between 1 and 24 media. Media have between 1 and 1,577 incoming flows (median = 1; average = 9.4). These flows connect them to between 1 and 285 distinct legal entities (median = 1; average = 3.2). 68 media maintain a continuous flow over all 15 quarters from between 1 and 18 legal entities.

The quantitative values of quarterly flows vary between € 5,000 (the minimum to be reported) and € 1.929.533 (median = € 10,931; average = € 23,444).

3.3 Design Requirements

Based on our data analysis described above and the interviews with data journalists interested in the MT dataset as reported by Niederer et al. [NRA+16], we can identify five design requirements that a visualization design for the MT dataset should fulfill:

R1 Data scalability: The number of vertices for both legal entities and media is relatively large. Besides the institutions' names and their network relations, there are no further data that could be used for clustering vertices. While a majority of vertices is only sparsely connected, some central vertices have a large number of flows. Likewise, the weights representing amount of money can vary widely within the network. The time dimension adds additional scale.

R2 Development over time: The data journalists interviewed by Niederer et al. [NRA+16] expressed particular interest in patterns or abnormalities in the number and weight of flows over time.

R3 Data wrangling: For two reasons, users would need to refine the MT dataset by basic data wrangling functionality: First, they can add their implicit expert knowledge into the analysis. For example, they could group together the federal ministries run by politicians of the same party.
Second, data quality is still not sufficient for some data entries even though data quality measures have been taken by the RTR and the dataset has been pre-cleaned by the Medientransparenz Austria project. Table 2 shows some examples based on media from this dataset containing the string "standard". It should be possible to combine entries with different forms of writing or different media (print, online, app) of the same newspaper and to hide entries of poor quality.

R4 Ease of use: The target audience of the MT dataset such as interested citizens or data journalists will most likely have no expert knowledge of statistics or visualization. They will access the MT visualization as a spontaneous activity where no special training can be provided. Therefore, care should be taken that well-known visualization techniques are chosen and the user interface is self-explaining.

R5 Interactive exploration: Some users will approach the MT visualization trying to verify an existing hypothesis but we expect that a majority of usage session will consist of undirected exploration in search for patterns of interest. For this, interac-

Table 2: Media matching the query string "standard" ordered by the number of connected legal entities and showing the aggregated sum of transferred money. Three entries are different forms of writing the same website. The fifth entry contains the names of six separate newspapers and stands as example of inconsistent data collection when he MT dataset was started in 2012.

medium	#rel.	summed flows
Der Standard	189	18,905,741 €
derstandard.at	64	2,768,875 €
www.derstandard.at	19	312,242 €
Der Standard KOMPAKT	2	44,745 €
Standard Verlagsgesellschaft m.b.H.	1	3,099,082 €
Krone, Kurier, Presse, Salzburger Nachrichten, Standard, Kleine Zeitung	1	90,874 €
derstandard.at App	1	11,510 €
ES Evening Standard Magazine	1	10,884 €
http://www.derstandard.at	1	9,938 €

tive features are needed that are usable and help users maintain overview.

4 Visualization Design

Based on these design requirements, we developed a visualization design for the MT dataset (Figure 2). This section describes the design and explains its underlying rationale. In Subsection 4.1 the individual diagram views of the design are presented. How the user is able to interact with them is described in Subsection 4.2 and how the views are linked with each other is delineated in Subsection 4.3.

4.1 Attribute Visualization

The MT dataset contains 5 data attributes. The columns of Table 1 display these data attributes. It is not possible to visualize every single data record of this table in the dashboard, therefore the records are aggregated and the aggregated information is displayed. [Mun14, p. 305]

For example Figure 2.A shows aggregated data of money transferred over time. For this, the data attribute "money amount" is summed for all data records with the same value in the data attribute "time". This reduces the data to only 2 data attributes and only 1 data record per quarter. The sum is a quantitative attribute and the quarters can be handled as ordinal attribute. A bar chart suites the task of looking up and comparing the values of the different quarters best [Mun14, p. 150]. A similar aggregation is visualized for sum of money transferred by legal background in Figure 2.B.

To visualize the distribution of a single quantitative attribute a histogram can be used [Mun14, p. 306]. Figure 2.C shows a histogram of the data attribute "money amount".

Figure 2.D is a histogram of the data attribute "trend". This attribute is derived from all amounts of money e_i flowing from a legal entity to a medium over time i. The trend T quantifies the relative difference of money transferred between the first half of the quarters $|Q_m|$ and the second half.

$$|Q_m| = \left\lfloor \frac{|Q|}{2} \right\rfloor \quad (1)$$

$$T = \frac{\sum_{i=|Q_m|}^{|Q|} e_i - \sum_{i=1}^{|Q_m|-1} e_i}{\sum_{i=1}^{|Q|} e_i} \quad (2)$$

The categorical data attributes legal entity and medium both have a large number of categories (see Subsection 3.2), which are too many to visualize them in a bar chart. Neither is it possible to aggregate the categories in a reasonable way. But the entries can be sorted by another aggregated quantitative data attribute so that only the most relevant ones are displayed For example, it is possible to sort legal entities by the sum of money transferred from them to media. Figure 2.E shows the details for the first 10 sorted legal entities as a table with 4 columns. The first column shows the name of the legal entity. The second column shows the sum of the transferred money to various media over time. Additionally a sparkline sized bar chart represents the distribution of the transferred money over time. This enables the user to see the trend over time [Tuf06, AMST11]. The third column displays the number of relations, i.e. the count of media receiving money from the legal entity. The forth column displays the "trend" as calculated by Equation 1. This data table visualization enables the user to receive detailed aggregated information for a few entities. Figure 2.F applies the same visual representation to the first 10 sorted media.

The last visualization in Figure 2.G shows the flow of money from legal entities to media using a chord diagram [KSB+09]. The aggregated amount of money is encoded with the length of an arc of a circle segment of the diagram. This allows the user to see from which legal entity how much money is transferred to which medium. Because there are too many different categories, placeholder segments are generated for legal entities and media, which contain all not displayed entries and aggregate the money for all of them.

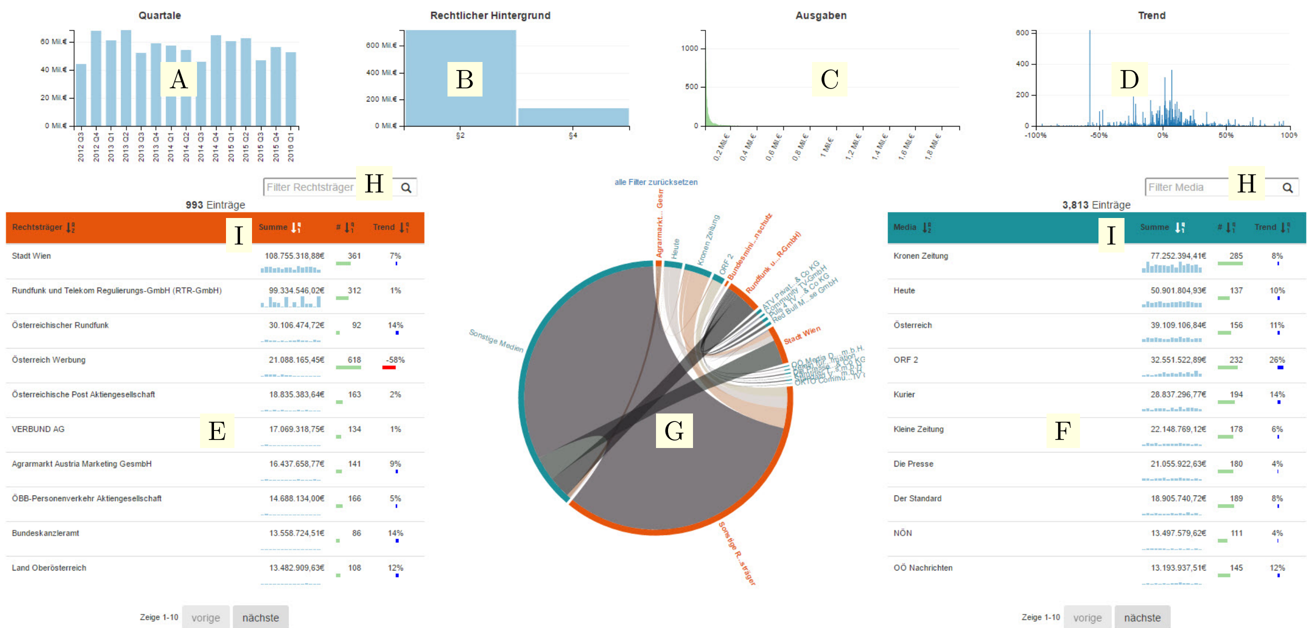


Figure 2: The visualization design for the media transparency (MT) dataset is comprised of seven views: (A) bar chart of aggregated money by time, (B) second bar chart for money by legal category, (C) histogram of money by flow, (D) histogram of increasing/decreasing trend, (E) table of 10 legal entities with total money, sparkline of money, number of connected media, and trend, (F) second table of top 10 media, and (G) chord diagram of flows. Both tables can be (H) searched and (I) sorted.

4.2 Interaction Components

The interaction with the diagrams is essential for the user to explore the data and to verify or refute an initial hypothesis (R5).

In Figure 2 the data is visualized without any manipulations by the user. The first three diagrams (Figure 2.A,B,C&G) give the user an overview of the underlying data. To analyze the data further the user is able to manipulate the view of the data.

Details on Demand The visualization design enables the user to receive details of a visual encoding of an aggregation of a data attribute of a chart. The visual encoding of a number, for example the height of a bar in a bar chart or the length of an arc in a flow visualization, supports the user to compare the encoding with the same encoded data attributes. To receive exact numbers the user is able to hover over each visual encoded element and receive in place information with a tool-tip [Dix09].

Select Elements To receive even more detailed information about the highlighted visual element, the user is able to select it. The data is then filtered by the selection and all other visualizations are updated with the newly filtered data. This interaction method is implemented in the visualization design as simple left-mouse-click and works for every visualization. By clicking onto a data table row the entity is selected. The histograms (Figure 2.C&D) do not support a simple click operation, but a click and drag operation to select a one dimensional range of the attribute in the histogram [Mun14, ch. 11.4].

Highlight Elements To visualize which visual elements are selected, the color saturation of the visual element is increased and for the filtered elements the saturation is decreased. Figure 3 shows this difference in saturation in contrast to the not selected visual elements in Figure 2.A&B. The coloring of the small-multiple bar charts in the data table is also linked with the highlighting of the time bar chart. The used colors are selected using the ColorBrewer2 tool, which is based on evaluation of "385 unique colour schemes [...] across different computer platforms and monitors, [...] for possible colour-blind confusions, as well as in printed formats." [HB03]

Sort To explore detail information for the trend over time, money, and the number of relations from one entity to another, the user is able to sort the data table along the data attribute of her/his interest (see Figure 2.I).

Search To support users' who want to analyze the data for a specific entity, full-text search is integrated. In our visual design this is implemented as a simple form text fields for the legal entities and media (see Figure 2.H).

Combine and Remove Like already mentioned in Section 3.1 the data quality might not be optimal. As modifying the underlying data cannot be expected by the target user group, interactive visual editing should be possible (R3). In our prototype, users may remove entries and combine multiple entries into a single entry. With a click onto the labels above a data table the selected rows of that table are combined or removed.

4.3 Coordinating Multiple Views

The designed interface connects the different visualizations and widgets and organizes them. The views are arranged on fixed positions, but the user is able to filter the data [EB11, Rob07]. Because all visualizations of the media transparency database use the same data set it is possible to link the selection between all views and thus use each view for dynamic query [AS94, ST98]. Additionally the color of the visual elements indicate which aggregation is used. This helps the user to see the connection between the visualizations and it enables the user to understand the connection of a data attribute in one to the distribution of a data attribute in another diagram [Mun14, ch. 12].

5 Implementation

The visualization design has been implemented as a web-based software using JavaScript with the libraries D3.js [BOH11], Crossfilter [cro], and dc.js Dimensional Charting [dc].

The implementation is available from `https://github.com/VALIDproject/mtdb2` as free and open source software under a BSD-2-clause license and can be tested at `http://medientransparenz.validproject.at/dashboard/`. For iterative refinement an informal usability test with two subjects was conducted.

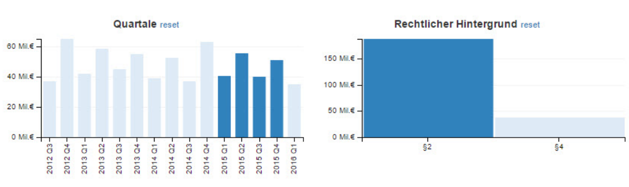


Figure 3: The elements of the visualizations adapt color saturation upon selection changes.

6 Usage Scenario

This sections presents a usage scenario to understand how the visualization of the MT dataset enables users to obtain a deeper insight into the data. The steps of the scenario can be followed in a video located at `https://vimeo.com/188278798`.

In most cases a user that is interested in a data set has some a-priori knowledge and a hypothesis that she/he wants to verify or falsify. In this scenario the user is interested in which legal entities spend money on online advertisement with Google.

Entering “google” into the full text search, the list returns 57 entries (e.g.: google, google.at, www.google.at, ...) due to data quality issues. By interactively manipulating the data the user is able to obtain a deeper insight. For example by combining the 57 categorical entries of the data attribute media to one entry named “Google”. The flow visualization is now easier to read because the number of visual elements was reduced.

The user is able to filter data which she/he is not interested in to obtain new information. For example by selecting only the entries of universities. This results in 3 legal entities, which the user is now able to compare in more detail (see Figure 4).

7 Conclusions

This paper presented a visualization design to explore the MT dataset, a large open government data asset reporting on the flows of money from government to media. We implemented the design as a web-based prototype, made it publicly available, and showcased it on science communications events like Lange Nacht der Forschung. Based upon these experiences and informal feedback we received, we can now reflect how well the visualization design addresses its design requirements and provide directions for future research:

R1 Data scalability: The various views of aggregated attributes are useful to provide a big-picture overview of the dataset. Subsequently, the interaction concept of linked selection, sorting tables, and showing the first results works to learn about the details. Some users criticized the chord diagram as being too cluttered and hard to read. A Sankey diagram is being considered as alternative. In future work, two additional proposals from the preceding problem characterization study by Niederer et al. [NRA+16] can be adopted: The large number of legal entities and media could be automatically clustered into hierarchical groups using text or network analytics. Alternatively, supplementary data could be loaded to provide additional properties such as geographic area for legal entities and/or media. These properties

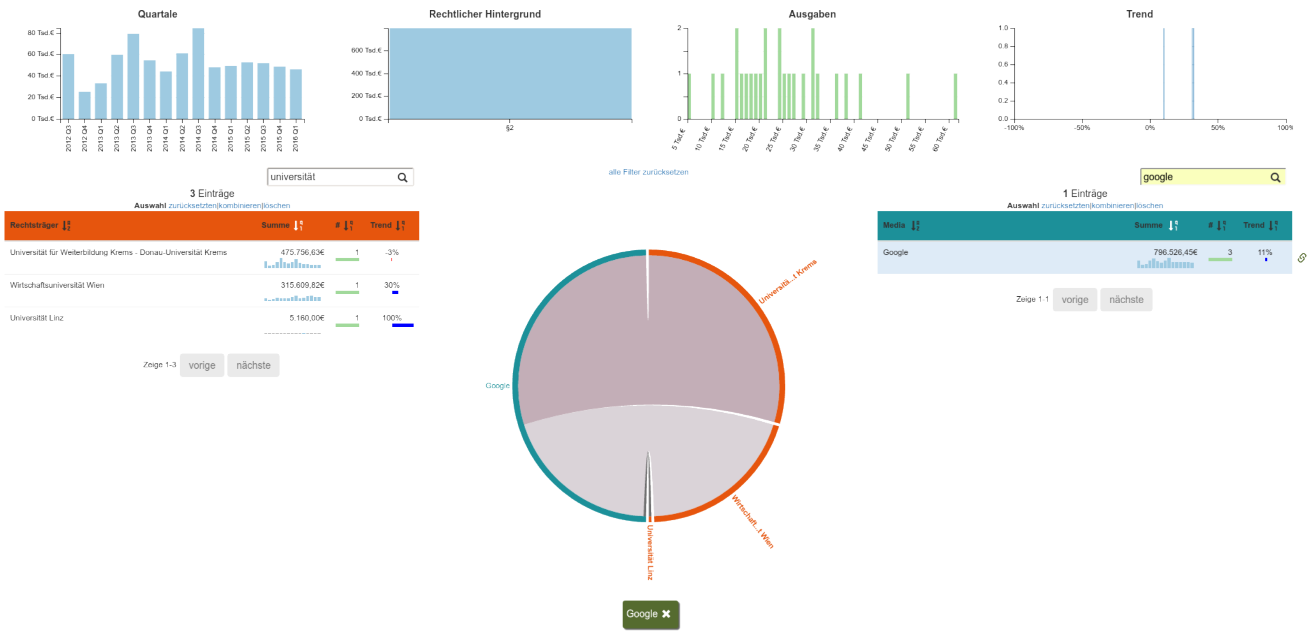


Figure 4: Snapshot of the visualization after following the steps of the usage scenario.

would subsequently be used for filtering and aggregation.

R2 Development over time: Both the bar chart view showing aggregated money flow over time, and the sparkline sized bar chart for each legal entity/medium work well to show distribution, abnormalities and other temporal patterns for the currently selected respectively visible items. The derived attribute "trend" was added to allow overview and direct manipulation of one concrete temporal pattern. While being a powerful feature, it is hard to grasp for novice users of the MT dataset visualization. Further design experiments are necessary to provide user-friendly exploration of temporal dynamic flows in bipartite networks.

R3 Data wrangling: The interactions to combine legal entities and/or media offers some benefits. The views are less cluttered by different entries for related institutions. In some cases data wrangling can eliminate a perceived false patterns such as abrupt end of flow to one medium that is in fact continued to a medium of a slightly different name.

Further work on data wrangling is indicated: On the one hand, we found the current functionality too limiting in several exploration sessions and desired more flexibility such as hierarchical groups and/or multi-group assignment like tags. On the other hand, the two functions 'combine' and 'remove' introduced more confusion for first time users. Possibly, a dedicated data wrangling mode should be provided so that these features are not visible by default.

R4 Ease of use: The visualization design is built using simple visual representation techniques that are well known to the general public. Still, the multiple views in composition were described as slightly overwhelming at first impression. In addition, novice users were not aware of direct manipulation so they did not expect they could filter the data e.g. by clicking on a bar.

R5 Interactive exploration: As demonstrated in the usage scenario, the visualization design allows free exploration of the MT dataset. While doing so, users can maintain overview of system state, i.e. which selections are active and also reset selections.

As further support for exploration, the data journalists interviewed by Niederer et al. [NRA+16] suggested documentation of the research path in order to provide analytic provenance [NCE+11].

Thus, our design study yielded not only a possible visualization design but also a range of directions for future work on exploring flows in dynamic bipartite networks.

Acknowledgements

This work was supported by the Austrian Ministry for Transport, Innovation and Technology (BMVIT) under the ICT of the future program via the VALiD project (no. 845598) and by the Austrian Science Fund (FWF) via the KAVA-Time project (no. P25489).

References

[AMST11] Wolfgang Aigner, Silvia Miksch, Heidrun Schumann, and Christian Tominski. *Visualization of Time-Oriented Data.* Springer, London, 2011.

[AS94] Christopher Ahlberg and Ben Shneiderman. Visual information seeking: Tight coupling of dynamic query filters with starfield displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence*, pages 313–317. ACM, 1994.

[Aus15] Julian Ausserhofer. "Die Methode liegt im Code": Routinen und digitale Methoden im Datenjournalismus. In Axel Maireder, Julian Ausserhofer, Christina Schumann, and Monika Taddicken, editors, *Digitale Methoden in der Kommunikationswissenschaft*, Digital Communication Research, pages 87–111. Berlin, 2015.

[BBDW16] Fabian Beck, Michael Burch, Stephan Diehl, and Daniel Weiskopf. A taxonomy and survey of dynamic graph visualization. *Computer Graphics Forum*, published online before print: 25 January, 2016. doi: 10.1111/cgf.12791.

[BKH+16] Benjamin Bach, Natalie Kerracher, Kyle Wm. Hall, Sheelagh Carpendale, Jessie Kennedy, and Nathalie Henry Riche. Telling stories about dynamic networks with graph comics. In *Proc. CHI Conf. Human Factors in Computing Systems*, pages 3670–3682. ACM, 2016.

[BOH11] M. Bostock, V. Ogievetsky, and Jeffrey Heer. D3: Data-driven documents. *IEEE Trans. Visualization and Computer Graphics*, 17(12):2301–2309, 2011.

[CMS99] Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman, editors. *Readings in Information Visualization: Using Vision to Think.* Morgan Kaufmann, San Francisco, 1999.

[cro] Crossfilter – fast multidimensional filtering for coordinated views. `http://square.github.io/crossfilter/`. Accessed: 2015-11-16.

[dc] dc.js – dimensional charting Javascript library. `https://dc-js.github.io/dc.js/`. Accessed: 2015-11-16.

[Dix09] Alan Dix. *Human-Computer Interaction.* Springer, 2009.

[EB11] Micheline Elias and Anastasia Bezerianos. Exploration views: understanding dashboard creation and customization for visualization novices. In *Proc. 13th IFIP TC 13 Int. Conf. Human-Computer Interaction, INTERACT*, pages 274–291. Springer, 2011.

[GBD09] Martin Greilich, Michael Burch, and Stephan Diehl. Visualizing the evolution of compound digraphs with TimeArcTrees. *Computer Graphics Forum*, 28(3):975–982, 2009.

[Ham] Markus Hametner. Inserate: 40,6 Millionen im ersten Quartal. `http://derstandard.at/2000017464403/`. WebStandard: 2015-06-15.

[HB03] Mark Harrower and Cynthia A Brewer. ColorBrewer.org: An online tool for selecting colour schemes for maps. *The Cartographic Journal*, 40(1):27–37, 2003.

[HBW14] M. Hlawatsch, M. Burch, and D. Weiskopf. Visual adjacency lists for dynamic graphs. *IEEE Trans. Visualization and Computer Graphics*, 20(11):1590–1603, 2014.

[HSS15] Steffen Hadlak, Heidrun Schumann, and Hans-Jörg Schulz. A survey of multi-faceted graph visualization. In Rita Borgo, F. Ganovelli, and Ivan Viola, editors, *Proc. Eurographics Conf. Visualization – State of The Art Report, EuroVis STAR*, pages 1–20. Eurographics, 2015.

[KSB+09] Martin Krzywinski, Jacqueline Schein, İnanç Birol, Joseph Connors, Randy Gascoyne, Doug Horsman, Steven J. Jones, and Marco A. Marra. Circos: An information aesthetic for comparative genomics. *Genome Research*, 19(9):1639–1645, 2009.

[Lan] Fabian Lang. Medientransparenz - die zweite. `http://www.paroli-magazin.at/555/`. Paroli-Magazin: 2013-03-18.

[Lor10] Mirko Lorenz. Status and outlook for data-driven journalism. In *Data-driven* *journalism: what is there to learn?*, pages 8–17. European Journalism Centre, 2010.

[Med15] Medienkooperations- und -förderungs-Transparenzgesetz, MedKF-TG. Austrian Legal Information System, 2015. https://www.ris.bka.gv.at/Dokumente/Erv/ERV_2011_1_125/ERV_2011_1_125.html. Accessed: 2016-08-31.

[Mun14] Tamara Munzner. *Visualization Analysis and Design*. CRC Press, 2014.

[NAR15] Christina Niederer, Wolfgang Aigner, and Alexander Rind. Survey on visualizing dynamic, weighted, and directed graphs in the context of data-driven journalism. In Hans-Jörg Schulz, Bodo Urban, and Uwe Freiherr von Lukas, editors, *Proc. Int. Summer School on Visual Computing*, pages 49–58. Frauenhofer Verlag, 2015.

[NCE+11] Chris North, Remco Chang, Alex Endert, Wenwen Dou, Richard May, Bill Pike, and Glenn Fink. Analytic provenance: process+interaction+insight. In *Proc. 2011 Ann. Conf. Ext. Abstracts Human Factors in Computing Systems, CHI EA*, pages 33–36. ACM, 2011.

[NRA+16] Christina Niederer, Alexander Rind, Wolfgang Aigner, Julian Ausserhofer, Robert Gutounig, and Michael Sedlmaier. Visual exploration of media transparency for data journalists: Problem characterization and abstraction. In *Proc. 10th Forschungsforum der österreichischen Fachhochschulen*. FH des BFI Wien, 2016.

[Rob07] Jonathan C. Roberts. State of the art: Coordinated & multiple views in exploratory visualization. In *Proc. Conf. Coordinated and Multiple Views in Exploratory Visualization, CMV*, pages 61–71, 2007.

[RTR] Austrian regulatory authority for broadcasting and telecommunications. Bekanntgegebene Daten. https://www.rtr.at/de/m/veroeffentl_medkftg_daten. Accessed: 2016-08-31.

[RTR16] RTR GmbH. Katalog Medientransparenz. Offene Daten Österreich, 2016. https://www.data.gv.at/katalog/dataset/58e02823-2bd2-4db7-9e2f-72a9ea7c7ffd. Accessed: 2016-08-31.

[SBSV] Peter Salhofer, Amir Basyouni, Mercedes Stibler, and Stephan Vrecer. Medientransparenz Austria. http://www.medien-transparenz.at/. Accessed: 2016-09-09.

[SMM12] Michael Sedlmair, Miriah Meyer, and Tamara Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE Trans. Visualization and Computer Graphics*, 18(12):2431–2440, 2012.

[ST98] Robert Spence and Lisa Tweedie. The Attribute Explorer: Information synthesis via exploration. *Interacting with Computers*, 11(2):137–146, 1998.

[Tuf06] Edward Rolfe Tufte. *Beautiful Evidence*. Graphics Press, Cheshire, CT, 2006.

[vdEvW14] Stef van den Elzen and Jarke J. van Wijk. Multivariate network exploration and presentation: From detail to overview via selections and aggregations. *IEEE Trans. Visualization and Computer Graphics*, 20(12):2310–2319, 2014.

[vLKS+11] Tatiana von Landesberger, Arjan Kuijper, Tobias Schreck, Jörn Kohlhammer, Jarke J. van Wijk, Jean-Daniel Fekete, and Dieter W. Fellner. Visual analysis of large graphs: State-of-the-art and future research challenges. *Computer Graphics Forum*, 30(6):1719–1749, 2011.

[WPZ+16] Yanhong Wu, Naveen Pitipornvivat, Jian Zhao, Sixiao Yang, Guowei Huang, and Huamin Qu. egoSlider: Visual analysis of egocentric network evolution. *IEEE Trans. Visualization and Computer Graphics*, 22(1):260–269, 2016.

[ZGC+16] Jian Zhao, Michael Glueck, Fanny Chevalier, Yanhong Wu, and Azam Khan. Egocentric analysis of dynamic networks with EgoLines. In *Proc. CHI Conf. Human Factors in Computing Systems*, pages 5003–5014. ACM, 2016.

A Review of Information Visualization Approaches and Interfaces to Digital Cultural Heritage Collections

Florian Windhager
Danube University Krems
Krems, Austria
florian.windhager@donau-uni.ac.at

Paolo Federico
University of Technology
Vienna, Austria
federico@ifs.tuwien.ac.at

Eva Mayr
Danube University Krems
Krems, Austria
eva.mayr@donau-uni.ac.at

Günther Schreder
Danube University Krems
Krems, Austria
guenther.schreder@donau-uni.ac.at

Michael Smuc
Danube University Krems
Krems, Austria
michael.smuc@donau-uni.ac.at

Abstract

After decades of digitization, the web hosts a large scale museum, consisting of millions of digital cultural objects. To balance the drawbacks of parsimonious search-centric interfaces, various approaches have been developed to enable also visual access to these collections, and to browse and explore the cultural richness of existing archives. This paper reviews information visualization approaches to digital cultural heritage collections, reflects on prominent arrangement principles and design choices for digital collection interfaces, and points out options for future research.

1. Introduction

From things making them smart (like tools, achievements, or information artifacts), to things lifting them up (art and entertainment) – cultures collect things. To share and preserve them for future generations, populations draw artful or useful objects (like texts, images, material objects, concepts, music, or films) together. These cultural heritage (CH) collections (libraries, galleries, museums, archives) contain notable works and objects – as well as associated knowledge and data.

With developing media technologies and collaborations, large digital meta collections (e.g. http://www.europeana.eu, http://trove.nla.gov.au/, or http://dp.la/) have emerged, which aggregate cultural heritage objects across institutions, domains, and countries, and make the web the largest museum ever around. Yet the situation is known to be rather bleak, when it comes to actually accessing the collected riches – not only, but especially for non-expert users, who often have no idea what to expect in the digital collection. The rampant problems with the widely dominant search box approach to cultural object collections have been thoroughly exposed and discussed [BOP82, DCW11, THC12, Whi15]. Whitelaw retells the typical search-based visit to online collections as a bizarre purchase order situation, where the widely dominant information retrieval paradigm over-successfully reduces data complexity (which in the CH context is often appreciated as its own reward), thus throwing the baby out with the bath water. Rather than throwing the collection doors open and offering multiple ways of access, visitors have to enter a drab (search box) lobby, which asks them "*yes, what?*" – and urges them to come up with demands towards the unknown [Whi15].

In contrast, more generous interfaces open up the digital archives' walls, tear down the drab lobbies, and offer multiple ways in, where they foster free-roaming, browsing and exploring, and support rich, serendipitous discoveries [DCW11]. We build on the multiply proven assumption, that information visualization (InfoVis) methods and techniques can strongly support such generous approaches. Yet according to our best knowledge, no systematic collection of InfoVis approaches to CH collections has been undertaken until now. To close this gap, we review related work and outline a possible classification of InfoVis approaches and interfaces for digital CH collections, which aims to consolidate the growing research field and to inform future projects.

In: W. Aigner, G. Schmiedl, K. Blumenstein, M. Zeppelzauer (eds.): Proceedings of the 9th Forum Media Technology 2016, St. Pölten, Austria, 24-11-2016, published at http://ceur-ws.org

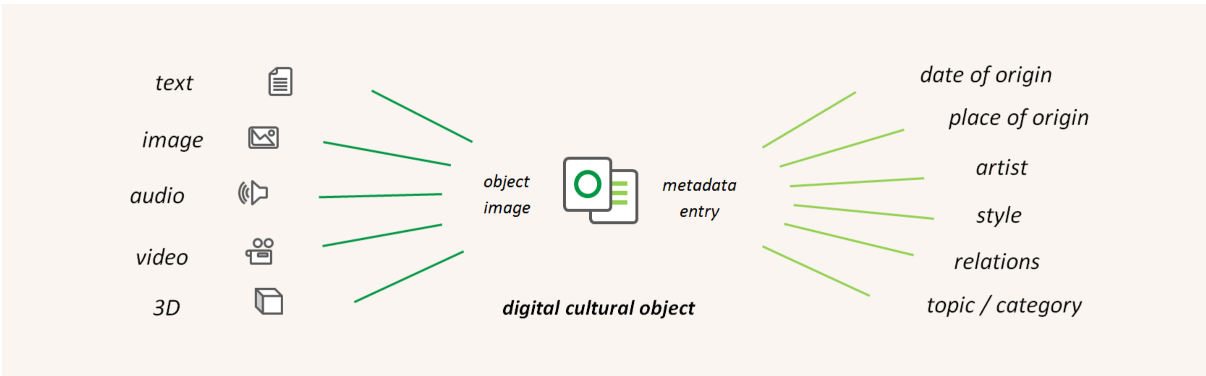


Figure 1: Common cultural object types (left) and common dimensions of object metadata (right).

2 Design Patterns for Interfaces to Digital Object Collections

If museums, libraries, or archives are the original three-dimensional display spaces for cultural object collections, their spatial arrangements are generated by a minimum of standard layouts: parallel tableaus on museum floors or in showcases, and linear arrangements along walls or shelves, ordered mostly due to the metadata dimensions of date, style, artist, or place of origin. Procedures of digitization extend cultural collections (complementing physical objects with digital ones) – and put their visual arrangement on digital display spaces up for renegotiation. For that purpose, all available metadata dimensions could be utilized – and furthermore encoded into novel collection representations.

Figure 1 illustrates the multitude of possible digital object types (left), and a selection of prominent metadata dimensions (right), with the latter being usually formatted due to a given documentation standard [Bac02]. This two-sided representation also mirrors the common dual nature of digital ob jects, duplicating an object into a realistic image of the object (provided by a spatial layout-preserving *scientific visualization* procedure), and a (semi)structured, multi-dimensional metadata entry. While the realistic image allows to study cultural objects in a close up-perspective, their accessibility in a larger collection is either provided by a search functionality – or by alternate, more generous approaches to interface design, including a wide variety of InfoVis images and methods. With interfaces thus taking over the role of museums or exhibition halls, their design determines an online collections' accessibility and impact, and should not be underestimated as a major factor for the overall success of any arts and culture mediation initiative.[1] We focus on the question how to visualize collection overviews and assemble relevant design patterns in the following sections, which will provide the categories for a more systematic recollection of InfoVis interfaces further down.

2.1 Close-ups, Previews and Collection Overviews

Cultural object collections commonly contain much more objects than could be displayed in a parallel close-up perspective on a screen. This challenge is commonly taken on by the design of macroscopic collection overviews – and their connection to vertical drill down and horizontal browsing options on demand [DCW11, GMPS00].

As a review of interfaces shows, collection overviews are usually following one of three design options: Whole object collections could be represented as i) multitudes of miniature previews (thumbnails), or

[1] Well knowing that the remote exploration of cultural collections on screens still "doesn't compare to being there" [RHQ14], digital interfaces mostly strive to augment and enrich traditional *in situ*-interaction with collections. This includes the design of approaches i) to provide macroscopic perspectives on high-volume collections in which patterns and relations become visible, ii) to extend visitors' working memory to grasp large, complex datasets often for the first time, iii) to add to richer, contextualized observations through linked data dimensions, or iv) to reduce collectors' and curators' biases and to facilitate more inclusive representations, suited for a broader user group [Sul13, GMD15].

as ii) multitudes of abstracted visual marks only (e.g. dots representing objects), whose arrangement principles are laid out in section 2.2. As a third option (iii), overviews can abstract from displaying separated objects, but encode selected object attributes into the visual variables of various diagrams (cf. 2.2.5), which opens up the field for the use of a wide spectrum of InfoVis methods, that can support further collection exploration too.

From a user and interaction perspective, overviews feature as natural starting or entry points to a collection. They provide initial orientation, and commonly enable further operations of zooming, filtering, and browsing to study details and close-ups on demand. While these transitions between micro and macro perspectives pose a central challenge for interaction design, we turn to prominent arrangements for macroscopic overviews first. As mentioned above, this is where various dimensions of object metadata (like place of origin, date of origin, artist, topics, or styles) come into play.

Figure 2 shows prominent arrangement principles for collection overviews: While the center left column features traditional ways of (multi)linear aggregations, the center right column lists methods for the visual encoding of *spatial* (i.e. cross-sectional, non-temporal) metadata aspects. Here "spatial" not only refers to geographic aspects of metadata, but also to their distributions in algebraic or vector spaces.

2.2 Encoding of Spatial Data Dimensions

Following a distinction by Kerracher et al. [KKC14], we distinguish methods of encoding *spatial* data dimensions from encoding methods for *temporal* (i.e. longitudinal) data aspects, which we consider to play a crucial role for the omnipresent time-orientation of CH collection data. Distributed across both sides of this distinction, we refer to the most prominent traditional spatial arrangement principle of object collections as (multi)linear arrangements (2.2.1), which are also frequently chosen for digital collection interfaces.

2.2.1 Lists, Slideshows, Grids and Mosaics

Mirroring the sequential arrangements in physical exhibitions along walls or shelves, vertical *lists* or

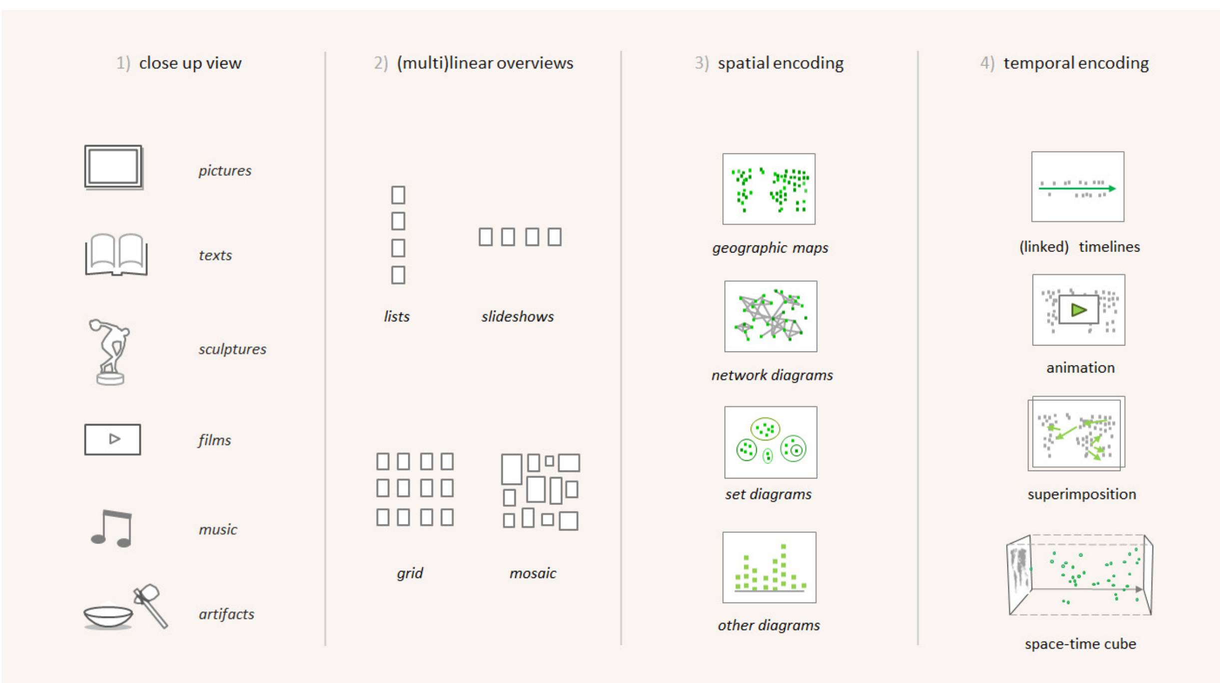


Figure 2: Principles for the visualization of cultural collections, from close-ups (left), to (multi-)linear aggregations (center left) to spatial (center right) and temporal (right) visual encoding methods.

horizontal *slideshows* arrange object collections in an unilinear sequence of previews on computer screens [IF:HTA]. As multilinear arrangements, *grids* and *mosaics* arrange previews in multiple rows, to raise the item-screen-ratio (Fig. 2, center left). In contrast to physical hangings, the guiding aspect for (multi)linear arrangements can often be freely chosen amongst existing metadata dimensions, so that either date of origin, alphabetical sequence, or even user metrics (like item popularity) determine the visible sequence of objects on screens [IF:GCI]. Furthermore, grids and mosaics can be dynamized, so that tiles represent object categories or subcollections and change their content over time, to enable also passive contemplation without clicking and scrolling [Whi15, para 39]. Going beyond (multi)linear arrangements, several InfoVis methods support the visual encoding and exploration of spatial (non-temporal) data aspects for whole collections.

2.2.2 Geographic Maps

As place of origin counts among the most frequently documented data dimensions of cultural objects and artifacts, *geographic maps* often serve as a visualization method to show the spatial distribution of artifacts' origins [BGSvdB14, IF:DGB, TO:GBDE, TO:PAL, TO:VS].

2.2.3 Network Diagrams

As for relational data (e.g. influences, references, inter-artifact relations) network diagrams allow users to explore the proximities and distances of artifacts or cultural actors in relational or topological spaces [HSC08, IF:DDBV, IF:ECB, IF:IA, IF:HG, IF:EDG, TO:PAL].

2.2.4 Set Diagrams

Given different thematic or stylistic classifications of cultural artifacts, set diagrams or treemaps offer insights into categorically and often also hierarchically structured object metadata constellations [XEJJ14, UPM12, IF:PAN].

2.2.4 Other Diagrams

When overviews abstract from single objects and focus on data distributions in different metadata dimensions, a wide variety of further InfoVis diagrams can provide overview on selected collection aspects, including area charts [IF:SCE], ring charts [IF:DDBV], scatter plots [Man09, ABO12, IF:CG], and many more.

These different diagrams again could be integrated into multiple coordinated views by CH collection dashboards [UTA10]. As an interesting crossover approach, diagrams could also be synthesized from object previews, allowing for seamless micro-macro transitions [IF:PVWF].

2.3 Encoding of Temporal Data Dimensions

While maps, networks, set and other diagrams provide specific insights into spatial data aspects and distributions, they initially offer static images for aggregated data only. Yet with *temporal aspects* (like date of origin) playing a crucial role in the domain of CH data, most interfaces have to encode temporal information too.

2.3.1 (Linked) Timelines

One prominent option is to represent time linearly, which is done with linear timelines as singular views, or with *linked timelines*, usually implemented as coordinated temporal view in addition to spatial representations [Kra16, IF:DGB, IF:HTA, IF:MOTW, IF:NL, IF:PAN, TO:VS].

2.3.2 Animation

Further options for encoding temporal data aspects build on the abovementioned spatial visualizations and add temporal information in a hybrid, spatiotemporal way. Among these, *animation* is frequently used, mapping time to time [IF:DGB, IF:PAN].

2.3.3 Superimposition

Superimposition approaches merge multiple temporal layers or snapshots into one visualization, with temporal data aspects often being distinguished by different colors [BGSvdB14], or visualization of movement trajectories [TO:NL].

2.3.4 Space-Time Cube

Space-time cube representations build on 2D planes of encoded spatial data dimensions (like maps or networks), and map time to an additional spatial dimension, i.e. the orthogonal z-axis. Cultural object collections thus arrange as characteristically shaped 3D point clouds, according to various spatio-temporal layouts [Kra05, WMS*16].

2.4 Multi-Method Interfaces

As the assembly of approaches and interfaces in *table 1* shows, multiple spatial encoding methods have already been implemented in the CH data domain – often also as multi-method interfaces to enable the combination of different exploratory views on the data. The same holds true for different temporal encoding methods: It is well known that different temporal encoding methods show different strengths and weaknesses. Due to this reason, advanced InfoVis interfaces increasingly combine multiple temporal and spatial encoding techniques, to compensate their drawbacks and add up their complementary benefits [KKC14]. This equals the provision of multiple access points and overviews [THC12], which form complementary composites, revealing different "parallax" views of a collection [Dru13]. Due to the relevance of this design principle, the following collection primarily takes approaches and interfaces into account which have been implementing a multi-method approach.

3 Assembling Information Visualization Approaches to Digital CH Collections

Table 1 provides an overview of prominent InfoVis approaches to digital CH collections. Interfaces are classified and specified according to four main categories. While the first two categories make the chosen spatial and temporal encoding methods visible (cf. 2.2 und 2.3), the third column specifies the focus of interest, which predominantly is either a certain type of cultural objects, or a focus on cultural actors (FCA), or a focus on cultural topics or styles. The fourth column points out whether the approach is of conceptual and prototypical nature, or whether it provides an open, web-based interface [IF:XYZ] or a

Table 1: Information visualization approaches and interfaces to digital cultural object collections, ordered according to their chosen method of spatial encoding, temporal encoding, entity focus, and type of project.

		spatial encoding					temporal encoding					focus on cultural objects or actors							type of project		
reference	DCHC visualization projects	lists, slideshows, grids, & mosaics	geographic map	network diagrams	set diagrams (categori	other diagram	(linked) timeline	animation	superimposition + cc	3D	other encoding	texts	images	artifacts	music / audio	film / video	artists / persons	topics / styles / areas	concept / prototype	interface	tool
[IF:GCI]	Google Cultural Institute	x					x						x				x	x		x	
[Whi15]	Generous Interfaces (Case Studies)	x						x					x					x	x	x	
[ABO12]	ViewShare	x	x			x					x		x							x	
[IF:HG]	histograph	x		x								x	x			x	x			x	
[IF:CS]	CultureSampo		x	x	x	x						x	x	x	x	x			x	x	
[IF:DGB]	GeoBrowser / Europeana 4D		x				x	x	x			x	x	x	x	x	x			x	
[TO:PAL]	Palladio		x	x								x	x	x	x	x	x				x
[BGS+14]	GLAM Map		x						x			x							x		
[WMS+16]	PolyCube		x	x	x					x		x	x	x	x	x	x	x	x		
[IF:HTA]	Heilbrunn Timeline of Art History		x				x						x	x						x	
[IF:NL]	Neatline | Omeka		x				x	x	x			x	x	x			x			x	x
[IF:ROL]	Republic of Letters		x	x			x					x					x			x	
[IF:IA]	The Invention of Abstraction [FCA]			x					x								x			x	
[DBBV]	DBB visualisiert			x		x	x		x			x	x	x	x	x				x	
[HSC08]	EMDialog			x		x					x	x					x		x		
[IF:ECL]	eclap-Browser - Social Graph			x									x		x	x				x	
[IF:EDG]	Edgemaps [FCA]			x			x										x		x	x	
[XEJ*14]	Xu et al. 2014				x														x		
[IF:PAN]	Pantheon [FCA]		x		x	x	x										x			x	
[IF:CG]	Culturegraphy			x		x										x				x	
[IF:KB]	Kindred Britain [FCA]		x	x		x	x		x								x			x	
[IF:SCE]	SelfieCity Exploratory	x				x							x							x	
[IF:PVFW]	Past Visions by Frederick William IV					x	x						x							x	
[Man09]	Imageplot Suite [TO:IPS]					x							x								x
[IF:MOTW]	Museum of the World						x			x		x	x	x						x	

tool [TO:XYZ], with which external DH collection data could be visually explored [cf. Pos16].

3.1 Interpretation

While looking at single approaches helps to specify their implemented combination of methods, parsing of columns helps to explore the prominence of encoding methods or object types. With regard to the overall distribution, well-established InfoVis techniques can be identified, as well as structural holes, which might deserve closer attention by future interface design and research. Exemplarily, the distribution of temporal encoding methods shows a dominant use of (linked) timelines, which again are known to evoke split attention effects [AS05]. To reduce cognitive, load more spatio-temporally integrated encoding techniques like space time-cube representations could be tested.

3.2 Limitations

Aiming for the consolidation of the research field and for orientation of future approaches, we are still aware of two obvious limitations. As the interaction with 'cultural object collections' is investigated in multiple academic domains, the current review is far from exhaustive. Yet by highlighting and comparing recent works and developments, we hope to lay ground for a more systematic and critical discussion – as well as for their future enrichment and refinement.

Furthermore, we consider the chosen categories of classification to be relevant from an InfoVis methods perspective, but are aware of possible other foci of attention. As such we exemplarily consider interaction and navigation techniques to provide productive categories or further analysis, as well as a wide variety of 'humanistic' user experience and design principles [Dru13, DCW12, Whi15], which could help to shape the focus on relevant DH interface functions and features with even more precision.

4. Conclusions and Outlook

We presented a review of InfoVis approaches and interfaces to digital CH collections, and arranged existing work by the means of a categorical framework, which we submit for critical examination and collective refinement.

We expect the field of CH collection visualization to further develop and diversify – not least due to the fact that the world wide web renders itself ever more indispensable as a medium for knowledge communication. Despite restricted budgets of local collectors and institutions, efforts for digitization and dissemination will continue, as will the development of web-based interfaces.

From an InfoVis perspective, we consider the field of CH data, users and tasks, to be a specifically productive one, revolving around grand design challenges. While CH data is often characterized by massively heterogeneous and time-oriented data complexity, its audiences approach it with heterogeneous, underspecified tasks [MFM*16]. Besides the consideration of well-known principles of graphical excellence, such casual users require also more aesthetics-oriented, entertaining approaches. In contrast to principles of parsimonious design and complexity minimization, the preservation of aesthetic complexity and diversity matter in the CH domain, and non-conclusive explorations provide their own reward. Therefore, the value of methods supporting horizontal browsing, multiple access points and serendipitous insight creation is ranging high. This makes CH data a challenging research field, expanding and enriching the scope of consolidated playing fields for InfoVis research far beyond expert-oriented professional applications.

Furthermore, we expect new options for interface design to emerge from the expansion and pervasion of *linked data* in the CH realm [KAR15, IF:CS], as well as the utilization of *user data,* which will open up new ways to weigh, highlight, recommend, and tailor interfaces for general audiences and specific user groups alike.

From a systematic point of view – which might be most relevant because of its *didactic* implications – we hope for a continued discussion and consolidation process to accompany the outlined developments. We consider such macroscopic reflections not only to be relevant for integrating the state of the art on academic grounds (informing new directions and approaches), but also for introducing visitors to the workings of their new online museums and archives. In contrast to traditional encounters with culture collections, their experiences and learnings will also depend on their ability to comprehend and master the powerful (re-) arrangement, encoding and interaction techniques, which new interfaces are already providing us with.

Acknowledgements

This research has been supported by the Austrian Science Fund (FWF), Project No. P28363

References

[ABO12] Algee, L., Bailey, J. Owens, T. (2012). Viewshare and the Kress Collection: Creating, sharing, and rapidly prototyping visual interfaces to cultural heritage collection data, *D-Lib Magazine 18*(3).

[Bac02] Baca, M. (2002). A picture is worth a thousand words: Metadata for art objects and their visual surrogates. *ALCTS Papers on Library Technical Services and Collections*, 131-138.

[BGSvdB14] A. Betti, D. Gerrits, B. Speckmann, & H. Van den Berg. (2014). GlamMap: Visualising Library Metadata. Proceedings of VALA.

[BOB82] Belkin, N.J., Oddy, R.N. & Brooks, H.M. (1982). ASK for information retrieval: Part I. Background and theory. *Journal of documentation 38*(2): 61–71.

[AS05] Ayres, P., & Sweller, J. (2005). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *Handbook of multimedia learning* (pp. 135-146). New York: Cambridge University Press.

[DCW11] Dörk, M., Carpendale, S., Williamson, C. (2011). The information flaneur: A fresh look at information seeking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1215–1224), ACM.

[Dru13] Drucker, J. (2013). Performative Materiality and Theoretical Approaches to Interface. *Digital Humanities Quarterly 7*(1).

[DWC15] Düring, M., Wieneke, L., & Croce, V. (2015). Interactive Networks for Digital Cultural Heritage Collections - Scoping the Future of HistoGraph. In P. Cimiano, et al. (Eds.), *Engineering the Web in the Big Data Era* (pp. 613–616). Springer International.

[GMD15] Glinka, K. Meier, S., Dörk, M. (2015). Visualising the »Un-seen«: Towards Critical Approaches and Strategies of Inclusion in Digital CH Interfaces, in: C. Busch, J. Sieck (Eds.), *Kultur Und Informatik (XIII) - Cross Media,* Hülsbusch, Berlin, pp. 105–118.

[GMPS00] Greene, S., Marchionini, G., Plaisant, C., & Shneiderman, B. (2000). Previews and overviews in digital libraries: Designing surrogates to support visual information seeking. *Journal of the American Society for Information Science, 51*(4), 380–393.

[HSC08] Hinrichs, U., Schmidt, H., & Carpendale, S. (2008). EMDialog: Bringing information visualization into the museum. *IEEE Transactions on Visualization and Computer Graphics, 14*(6), 1181–1188.

[KBM16] Kontiza, K, Bikakis, A., Miller, R. (2015). Cognitive-based Visualization of Semantically Structured Cultural Heritage Data, *Proceedings of the International Workshop on Visualizations and User Interfaces for Ontologies and Linked Data.* URL: http://ceur-ws.org/Vol-1456/

[Kra05] Kraak, M. J. (2005). Timelines, temporal resolution, temporal zoom and time geography. *In Proceedings of the 22nd International Cartographic Conference*. A Coruña Spain.

[Kra16] Kräutli, F. (2016). *Visualising cultural data: exploring digital collections through timeline visualisations*. Doctoral dissertation, Royal College of Art.

[KKC14] Kerracher, N. , Kennedy, J., & Chalmers, K. (2014). The design space of temporal graph visualisation. In *Proceedings of the 18th EuroVis* (Vol. Short). Swansea: Eurographics.

[Man09] Manovich, L. (2009). *Cultural Analytics: Visualizaing cultural patterns in the era of "more media".* Domus.

[MFM16] Mayr, E., Federico, P., Miksch, S., Schreder, G., Smuc, M., & Windhager, F. (2016). Visualization of Cultural Heritage Data for Casual Users. Proceedings of the 1st Workshop for Visualization for the Digital Humanities, Baltimore, MD.

[Pos16] Posner, M. (2016). *The Digital Art Historian's Toolkit*.URL:http://program.dh.ucla.edu/getty/index.php/the-digital-art-historians-toolkit/.

[RHQ14] Rogers, K., Hinrichs, U., & Quigley, A. (2014). It doesn't compare to being there: in-situ vs. remote exploration of museum collections. *The Search Is Over! Exploring Cultural Collections with Visualization*, London, UK. http://searchisover.org/

[Sul13] Sula, C. A. (2013). Quantifying Culture: Four Types of Value in Visualisation. In J. Bowen, S. Keene, & K. Ng (Eds.), *Electronic Visualisation in Arts and Culture* (pp. 25–37), Springer.

[THC12] Thudt, A., Hinrichs, U., & Carpendale, S. (2012). The bohemian bookshelf: supporting serendipitous book discoveries through information visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1461–1470). ACM.

[UTA10] Urban, R.J., Twidale, M.B., Adamczyk, P.D. (2010). *Cultural Heritage Information Dashboards*.URL:http://www.ideals.illinois.edu/handle/2142/14936

[Whi15] Whitelaw, M. (2015). Generous Interfaces for Digital Cultural Collections, *Digital Humanities Quarterly*, 9.

[WMS*16] Windhager, F. Mayr, E., Schreder, G., Smuc, M., Federico, P., & Miksch, S. (2016). Reframing Cultural Heritage Collections in a Visualization Framework of Space-Time Cubes. In M. Düring et al.(eds.) *Proceedings of the 3rd HistoInformatics Workshop*, (pp. 20–24), Krakow. http://ceur-ws.org/Vol-1632/

[XEJJ14] Xu, W., Esteva, M., Jain, S.D., & Jain, V. (2014). Interactive visualization for curatorial analysis of large digital collection, *Information Visualization 13*: 159–183.

CH InfoVis Web-Interfaces

[IF: CG] *Culturegraphy*
URL: http://www.culturegraphy.com/

[IF:CS] CultureSampo
URL: http://www.kulttuurisampo.fi/?lang=en

[IF:DDBV] *Deutsche Digitale Bibliothek Visualisiert*
URL: http://infovis.fh-potsdam.de/ddb/

[IF:DGB] *DARIAH-DE Geo-Browser / Europeana4D*
URL: https://geobrowser.de.dariah.eu/

[IF:ECB] *eclap-Browser / Social Graph*
URL: http://www.eclap.eu/portal/

[IF:EDG] *Edgemaps* [FCA]
URL: http://mariandoerk.de/edgemaps/demo/

[IF:GCI] *Google Cultural Institute*
URL:https://www.google.com/culturalinstitute

[IF:HG] *histograph*
URL: http://histograph.cvce.eu/

[IF:HTA] *Heilbrunn Timeline of Art History | The MET*
URL: http://www.metmuseum.org/toah/chronology/

[IF:KB] *Kindred Britain* [FCA]
URL: http://kindred.stanford.edu/

[IF:IA] *Inventing Abstraction 1910-1925* [FCA]
URL: www.moma.org/inventingabstraction

[IF:MOTW] *Museum of the World*
URL: https://britishmuseum.withgoogle.com/

[IF:PAN] *Pantheon* [FCA]
URL: http://pantheon.media.mit.edu/

[IF:PVFW] *Past Visions by Frederick William IV*.
URL: https://uclab.fh-potsdam.de/fw4/

[IF:ROL] Republic of Letters [FCA]
URL: http://ink.designhumanities.org/dalembert/

[IF:SCE] *SelfieExploratory | SelfieCity*
URL: http://selfiecity.net/selfiexploratory/

CH InfoVis Tools

[TO:GBDE] *Geo-Browser Datasheet Editor*
URL: https://geobrowser.de.dariah.eu/edit/

[TO:IPS] *ImagePlot Suite*
URL: http://lab.softwarestudies.com/p/software-for-digital-humanities.html

[TO:NL] *Neatline | Omeka*
URL: http://neatline.org/

[TO:PAL] *Palladio*
URL: http://hdlab.stanford.edu/palladio/#/

[TO:VS] *ViewShare*
URL: http://viewshare.org/

Visual Exploration of Hierarchical Data Using Degree-of-Interest Controlled by Eye-Tracking

Nelson Silva[1,2], Lin Shao[1], T. Schreck[1], Eva Eggeling[1,2], and Dieter W. Fellner[1,3]

[1]Institute for Computer Graphics and Knowledge Visualization, Graz University of Technology, Austria
[2]Fraunhofer Austria Research GmbH, Austria
[3]Interactive Graphics Systems Group, Technische Universitaet Darmstadt, and Fraunhofer IGD, Germany

Abstract

Effective visual exploration of large data sets is an important problem. A standard technique for mapping large data sets is to use hierarchical data representations (trees, or dendrograms) that users may navigate. If the data sets get large, so do the hierarchies, and effective methods for the navigation are required. Traditionally, users navigate visual representations using desktop interaction modalities, including mouse interaction. Motivated by recent availability of low-cost eye-tracker systems, we investigate application possibilities to use eye-tracking for controlling the visual-interactive data exploration process. We implemented a proof-of-concept system for visual exploration of hierarchic data, exemplified by scatter plot diagrams which are to be explored for grouping and similarity relationships. The exploration includes usage of degree-of-interest based distortion controlled by user attention read from eye-movement behavior. We present the basic elements of our system, and give an illustrative use case discussion, outlining the application possibilities. We also identify interesting future developments based on the given data views and captured eye-tracking information.

In: W. Aigner, G. Schmiedl, K. Blumenstein, M. Zeppelzauer (eds.): Proceedings of the 9th Forum Media Technology 2016, St. Pölten, Austria, 24-11-2016, published at http://ceur-ws.org

1 Introduction

In this paper, we consider using eye-tracking information to create an adaptive Visual Analytics system. A main idea of Visual Data Analysis is to support analytic reasoning by interactive visual interfaces to data. This typically involves the integration of capabilities of data analysis in terms of visual information exploration, and the computation capabilities of computers to create capable knowledge discovery environments [KMSZ06, ABM07]. The need for effective data analysis solutions is obvious as more and more digital information is being generated and collected in many areas, e.g., in medicine, science, education, or business. Data analysis problems can be diverse, such as the amount and speed of data being generated, while it needs to be processed and analyzed. Other problems are related with the filtering, aggregation and visualization of this same data.

There are different types of data, among which graphs and networks are important data structures to model many relevant data sets [vLKS+11, HMM00]. The sizes of graphs grow quickly in many domains, and these sizes hinder visual exploration of the data, as visualization of large graphs is a challenge. As the above cited surveys show, there are numerous graph visualization methods available, however displaying just a few thousands of nodes effectively remains a problem. Therefore, data reduction, e.g., by clustering/collapsing of graph nodes is a common approach to limit the data complexity. In terms of hierarchies (trees, or dendrograms) these can be reduced to show only a certain depth of the hierarchy, and group together all elements of a sub tree at the sub tree root.

Traditionally, we can use pointing devices to zoom, pan and navigate to other areas of the graph, expand/collapse nodes, or adjust the level of abstraction [GST13]. However, several problems may arise with the mouse navigation, zoom and expand collapse strategies. When the user pans a graph, the mouse click can be anywhere on the graph, including empty parts of it; also, when the user stops panning, the mouse cursor is "parked" somewhere in the visualization, and no useful information can be inferred regarding the user intents.

Existing *eye-tracking* devices allow to track the fixation areas of a user in front of a display. Among others, eye-trackers are often used for evaluation purposes, or to experimentally study human visual attention. In our work we are interested in the question, if eye-tracking information can benefit the visual data exploration process, in addition or as an alternative to standard interaction approaches. Nowadays, we can use affordable eye-tracker devices [SZCJ16], like the EyeTribe system[1] to monitor the behaviors of the users while exploring a visualization. According to our experience, the device allows a useful tracking of the user gazing on specific regions and individual nodes, on a comparably large tree to be explored.

We present a concept and preliminary implementation of an approach to apply eye-tracking for the purpose of supporting visual exploration of large graphs. Our assumption is that the user gaze indicates areas of interest in a tree, and consequently we can use this information to dynamically expand or compress parts of the tree. Furthermore, we can also capture a *visual history* of the exploration process during which a user explores a tree view, with applications for example of 'replaying' analysis sessions or documenting interesting findings done along the way. This paper is our first step towards an experimental system by which we can explore design alternatives for eye-based interaction and visualization, as well as to conduct user studies.

2 Related Work

We briefly provide an overview of possible applications of eye tracking in evaluation, and as an interaction modality. We also discuss visualization of hierarchic data and degree-of-interest techniques.

2.1 Eye-Tracking and Applications

Eye-movement tracking is a method that is used to study, among others, usability issues in Human Computer Interaction (HCI) contexts. Pool and Ball [PB05] give an introduction to the basics of eye-movement technology, and present key aspects and metrics of practical guidance in usability-evaluation studies for capturing user eye movements as an input mechanism to drive system interaction.

Etemadpour et al. [EOL14] address eye-movement tracking on user studies regarding the accomplishment of typical analysis tasks for projected multidimensional data, such as tasks that involve detecting and correlating clusters. The authors examine and draw conclusions on how layout techniques produce certain characteristics that change the visual attention pattern.

In lab-based user experiments using eye-movements tracking, large and complex gaze trajectory data sets are generated. There is work which develops tools to help understand eye-movement patterns [BKR+14]. These should support the definition and exploration of a large number of areas-of-interest (AOIs). Eye-movement tracking data is usually analyzed using different methods [HNA+11] and visualization techniques [BKR+14]. Our work follows an AOI-based approach.

One important question refers to the justification of why should one use eye-movement tracking and not just the typical pointing devices, such as, the mouse. Many past studies debate the correlation between eye-movement tracking and mouse movements. These studies presented values from as high as 84% (in a study from 2001) [CAS01], to 69% [Coo06], to as low as 32% [RFAS08] (in a study from 2008) of correlation between eye and mouse movements. These results are usually dependent on the design of the user interfaces.

Another relevant discussion is centered on the many advantages of using eye-movement tracking analysis and on how to perform a correct eye-movement tracking evaluation [JK03]. Eye-movement tracking allows for a fast and continuous tracking of the interest of the users in real time, allowing the detection of moments of confusion, indecision and high interest regions [GW03]. Also, previous studies discuss an important link between cognitive processes and eye-movements [Hay04]. The accuracy of eye-movement tracking can be kept high by designing a user interface where the size of the areas of interest is big enough and in accordance to the eye-tracker characteristics and the experiment setup. More and more, new eye-trackers are also less affected by negative technical factors, e.g., the users' head movements that usually reduce the accuracy of the eye-tracker device and calibration difficulties.

2.2 Hierarchy Visualization and Focus-and-Context

In this work we consider visualization of hierarchic data. While hierarchies arise in many contexts, one prominent use of hierarchies is in data clustering. Generally, hierarchical graphs together with one of the many clustering techniques [MRS08] can form beneficial tools for the visual exploration of large data sets.

[1] `https://theeyetribe.com` (accessed 09/2016)

This visual exploration can be done in the form of trees (or dendrograms), due to their potential for visual abstraction [CdART12]. In a hierarchical clustering, users may chose the level of detail by which they explore data. Areas more close to the root contain more aggregate information, and areas closer to the leafs include more detail data.

An effective interaction technique for navigating large visualization spaces is to control the level of detail information shown throughout a given visualization. Furnas [Fur86] defines a degree of interest function (DOI-function) where to each node in the graph structure an interest score is defined. This score in turn is used to expand important areas while reducing other less important areas. Lamping et al. [LR96] demonstrated a focus+context (fisheye) scheme for visualizing and manipulating large hierarchies. Generally, the expansion or reduction can operate on different aspects, e.g., on the geometric, semantic or data-oriented level. Previous work [PGB02] was done regarding the dynamic re-scaling of branches of the tree to best fit the available screen space with an optimized camera movement. Concerning the aspect of developing adaptive visualizations, an important survey [CK15] was presented that highlights many techniques for emotion-driven detection, measurement and adaptation, among others. These are very relevant for our concept of adaptation based on degree-of-interest.

3 Concept for Visual Exploration of Hierarchical Data Guided by Eye Tracking

Next, we introduce our concept (Figure 1) for exploring hierarchical data based on degree-of-interest, guided by eye tracking. We will exemplify our concept by using scatter plots as the leaf elements in the hierarchy, which are to be explored by a user. The hierarchy is created by a hierarchical clustering algorithm using feature-based similarity between scatter plots. Our approach relies on eye-tracking to determine the degree-of-interest, which in turn distorts (i.e., magnifies/compresses) the hierarchy display.

3.1 Considered Application: Hierarchy of Scatter Plots

For the exploration of complex data sets, target visualization techniques such as scatter plots, parallel coordinates or glyph representations can be used to discover interesting findings in the data. In our approach, we rely on a set of scatter plot visualizations to represent all pairwise combinations of a high-dimensional data set. To explore a potentially large set of scatter plots hierarchically, we apply hierarchic clustering. Input to the clustering is a distance matrix between the set of scatter plots. The latter is obtained making use of image features, which have been shown to work well for the comparison of scatter plots [SvLS12]. More precisely, we compute a 25-dimensional intensity histogram for each plot. Then, we use the Euclidean distance between histograms to compute the distance (average linkage) of each plot. Using these visual features, the scatter plots can be arranged hierarchically (e.g., in a tree or circular layout) and the exploration for visually similar plots becomes more efficient.

3.2 Hierarchical Layout of Scatter Plots

Typically, there are a large number of scatter plot views for a high-dimensional data set, these views grows quadratically with the number of data dimensions. Specifically, an n-dimensional numeric data set can be represented in $\frac{n\times(n-1)}{2}$ distinctive views using two distinctive dimensions. To facilitate the exploration, we take the computed feature vectors of the scatter plots and apply a hierarchical clustering to structure the plots based on their visual similarities. Thus, we receive a structured representation of the space of scatter plots that arranges similar scatter plots spatially close. To create the hierarchy structure, we compute the average distance (average linkage) of each scatter plot and build a dendrogram tree, which contains all scatter plots on the leaf node level, see Figure 3. As usual, the dendrogram height describes the similarity (histogram distance) of the scatter plots.

3.3 Degree of Interest for Navigation of the Hierarchy

The above described dendrogram provides a useful spatial organization of the input space (scatter plots). Yet, the tree may still be large and complex, especially if we have a large number of leaf and internal nodes to inspect and compare. Hence, we introduce *spatial distortion* to enlarge parts of the tree currently being looked at by the user, while visually aggregating the remainder of the tree. To this end, we apply eye tracking using an EyeTribe (see Section 1) setup to track user gazes. Specifically, we measure the user attention on the tree nodes to compute a degree-of-interest (DOI) score for the elements of the dendrogram. Initially, we show the overall dendrogram using semantic zoom to fit the whole hierarchy onto the display space. From there, the user starts the graph exploration from any point in the view space. While the user navigates through the view space, the eye tracker captures the gaze path. When the user explores specific branches of the tree or local nodes, the eye gaze path and eye-fixation durations are recorded for each link and node of the tree. Therefore, besides a measure of interest based on similarity between scatter plots, we can now update interest metrics based on time and number of visits to a node.

Potentially, this recording can also be done for local parts of the scatter plots, i.e., tracking if the user is dedicating more viewing time to certain local areas in a plot. Such analysis may be useful to detect e.g., correlations, dense areas or clusters in a given plot. Each scatter plot involves the representation of variables (for x and y axes respectively), the interest of the user on these variables (axis) can also be tracked. In the next section, we apply the current eye gaze location in order to focus the display using semantic zoom. Conceptually, more applications are possible (see also Section 5).

3.4 Degree-of-Interest Visualization Using Eye-Tracking

We apply eye-movement information to allow the user to navigate through a hierarchy of clusters of scatter plots using semantic zoom. We define an *eye-tracking mode*, which if enabled, controls the expansion and collapsing of sub-trees in the display based on eye fixation. Specifically, the area where the user looks at is visually expanded, revealing the scatter plots under the sub-tree. The neighboring (remaining) sub-trees are represented using just node and link symbols. While they do not show particular scatter plots, this reduced representation is still indicating basic data properties like number of scatter plots represented, or structure of the similarity relationships within the dendrogram.

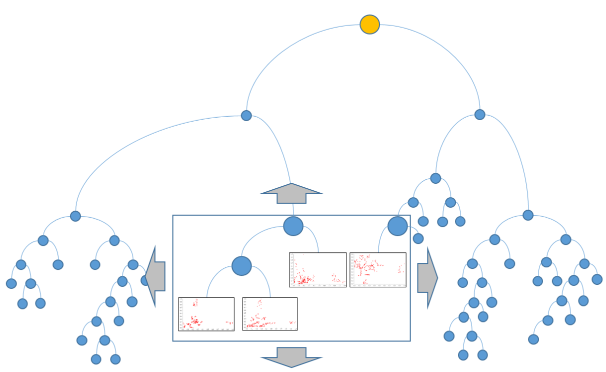


Figure 1: Concept: Exploration of large scatter plot spaces. The user eye-gaze is detected, leading to an expansion of the focused sub-tree (center rectangle). The remaining data is shown using a node-link representation (context, outside center rectangle).

When the user stops the *eye-tracking mode*, the application goes back to a state where it tracks only the user interest on each specific node, i.e. eye-gaze duration on each node (no pan control). We also show an overview of dendrogram areas visited so far (see Figure 2 for a gaze history view). This view allows to keep track of visited and unvisited areas, and constitutes input for further data analysis (see also Section 5).

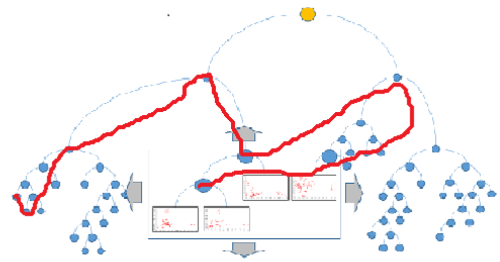


Figure 2: Gaze History Mock-up: It can be activated in the navigation panel. The user can track tree areas explored so far by an overlaid trace path (red line). It serves as a global map of explored/unexplored areas, and it is used for further analysis (see Section 5).

3.5 Benefits of Our Approach

We did informal, preliminary tests of our proposed navigation with 10 users. The feedback so far was positive, both to the semantic zoom mode and the gaze history view. The navigation was considered as rather smooth, and users can navigate without larger difficulties. Just by looking slightly away in the tree view, the corresponding movement is initiated in a very intuitive way. This facilitates the entire process of exploring the data in the tree, i.e., the view panning is synchronized with the field of view and the eye-movements of the user. When the user stops using the eye-based navigation, attention information starts to be collected again (eye-gaze duration on each area-of-interest) and it is the basis for the analysis of user interest detection and possible subsequent recommendation of interesting views. Note that in our concept we consider only gaze-based navigation. Of course, we can rely in addition on mouse/keyboard input to facilitate navigation, e.g., for labeling, saving views/bookmarking, etc.

4 Implementation and Application

To test our approach, we developed a proof-of-concept allowing the exploration of a large tree (dendrogram).

4.1 System Implementation

In our tree visualization, the leaf nodes are composed of visual representations of the data, i.e., scatter plots with pairs of data attributes. We use our own modified version of the JUNG system [OFWB03] for the tree visualization. We made changes on the adjustment of the lens size in the view space and positioning of the lens, now they are controlled by user eye-movements and updated in real-time. We created a customized tree layout to display color-coded nodes according to the computed similarity distance measure (darker color = lower similarity), and also the ability to display visual representations of the data on the leaf nodes, i.e, scatter plots.

The initial preset for DOI specification is the calculated similarity distance between scatter plot images. For this calculation, we make usage of a basic descriptor from the Java Image Processing Cookbook [2] that is based on the average calculation of 25 color triples for each image. After performing the comparison between each image using the descriptors, we create a distance matrix with the computed distances between all images. This matrix is handed out to an agglomerative hierarchical clustering algorithm [Beh16].

We tested our system with several hierarchical trees. Here, we illustrate the application of a hierarchical tree exploration of our data set. At the root and top subtrees we can find information about clusters of similar scatter plots, and at the leafs we find individual scatter plots. For this proof-of-concept we use a dendrogram comprising 269 scatter plots (leaf nodes) and 536 edges. Figure 3 shows a zoomed in view of the tree, the color-coded nodes according to distance similarity of the scatter plots, and the circular zooming lens (in gray color). In the navigation panel (top-right corner), we can get an overview of the entire tree size and respective available view space. The current view size (depending on the zoom level) and location is denoted by a white rectangle. The lens can be used to perform a close zoom into the scatter plot image, and it can be used to activate the display of a different visual representation of the data.

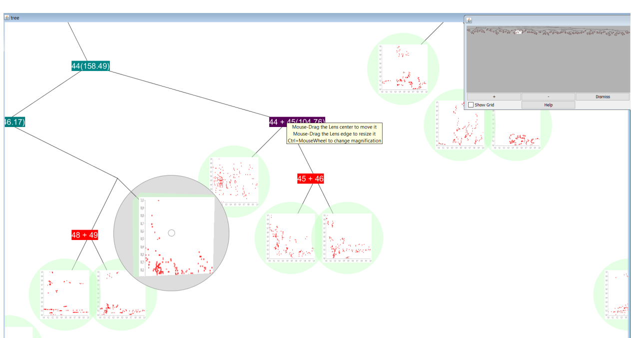


Figure 3: Zoomed-in view of the hierarchical tree. The mouse wheel can be used to: increase/decrease the lens magnifying ratio; increase/decrease the size of the circular zooming lens (gray circle) by clicking and dragging its border. A navigation panel (top-right corner) gives an overview of the actual position in the tree.

[2] Java Image Processing Cookbook (`http://goo.gl/FBXbjp`)

4.2 Application

Our data set is retrieved from the *Eurostat*[3] data repository, which provides a collection of data sets containing information on EU related topics (e.g., economy, population and industry). We use a preprocessed data set from preliminary work [SSB+15], which contains 27 statistical attributes from 28 EU countries showing temporal changes over time.

All navigation actions presented in the following examples illustrate a typical usage of our navigation system. Figure 4 shows an example of an ideal view over a small data portion, where the user is able to see the majority of the data. Practically, for larger data sets users will often have a more narrow view over the entire tree, depicted in Figure 5 with the 3 narrow views.

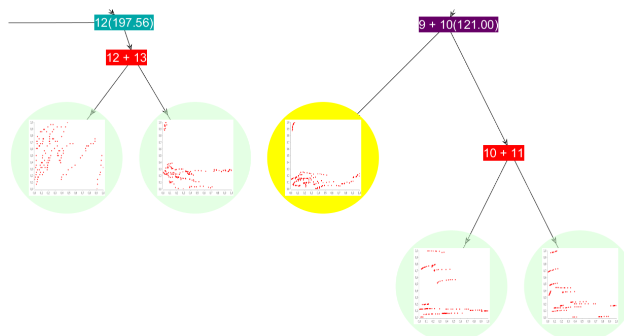


Figure 4: Ideal Case: Users can view several clusters of related scatter plots at once. Due to limitations of display space, this is often not possible, hence the need for adaptive visualization for navigation.

The navigation order (view sequence) followed by the user on these narrowed views can be random, it might just follow the similarity distance measures (depicted by the color-coded node rectangles). Figure 5, shows that the user first moved to view V1, where a group of interesting clustered scatter plots is visible (Figure 6). In view V1, the system detects a high gaze duration and infers that the interest is on one of the scatter plots (marked with "*"). After an in-depth inspection of this area, the user navigates to a view V2 (Figure 7) over another group of clustered scatter plots. The next most interesting and similar scatter plot (yellow color) is occluded in view V2 and it is only visible in view V3 (Figure 8).

It might take time until an interesting scatter plot (view V3) is spotted by the user. Also, there might exist other interesting scatter plots in another part of the tree, in a more far, and yet hidden location, e.g., Figure 9.

[3] Statistical Office of the European Union (`http://ec.europa.eu/eurostat`). Accessed 09/2016.

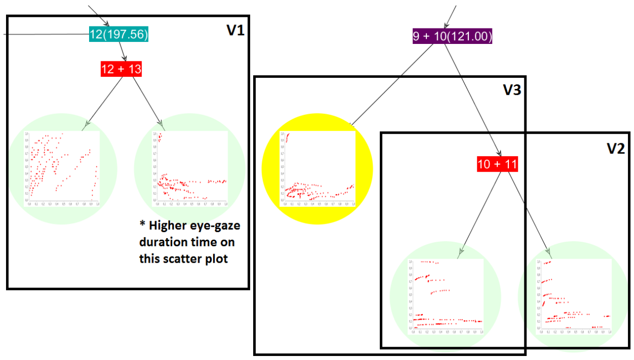


Figure 5: In practice, users may have limited views over a large space that must be explored. Therefore, the views (V1, V2, V3) might be limiting and not following an ideal sequence of exploration that would lead to finding interesting factors and to the creation of a useful mental model while exploring the data set. We take the eye-gaze duration time in account to infer about the interest of the user.

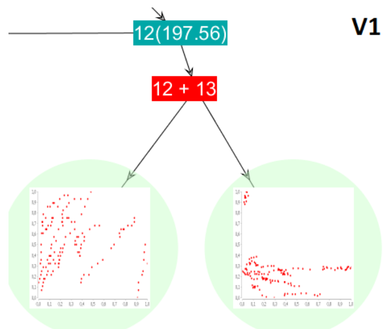


Figure 6: View 1: Realistic view of a first set of scatter plots. The user is focused on a set of scatter plots and unaware of other interesting locations of the tree.

In summary, our example merely demonstrates some of the challenges associated with the exploration of large graphs. The duration of the gazes can be used to expand or collapse sub-trees and hence provide a more organized (less cluttered) overview of the data, reducing the risk of getting lost in the exploration of large dendrogram trees. However, our measures computed directly from the location of the gazes are only a first step to control the views. We plan to collect data about the eye-movement scan paths and respective eye-gaze durations, as well as recurrences to develop more adaptive hierarchy views.

5 Discussion and Extensions

We implemented a proof-of-concept system for which we see numerous extension possibilities. First, our solution allows not only to adjust the amount of visual

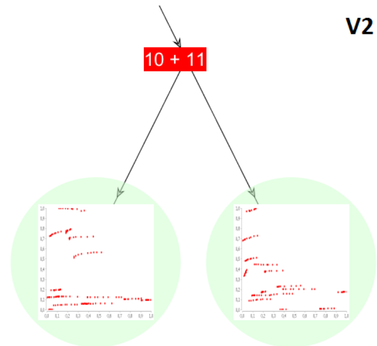


Figure 7: View 2: The user moves to a new location, but misses an interesting scatter plot that is occluded on a top location.

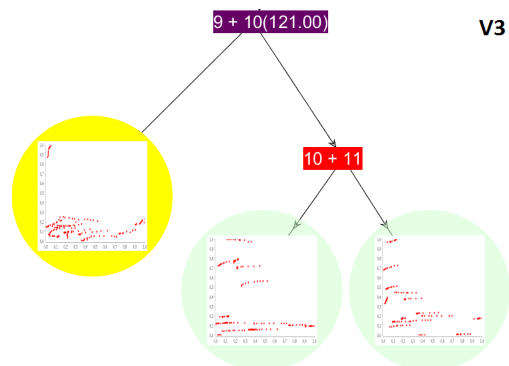


Figure 8: View 3: User navigates to this location and finds an interesting scatter plot related to view V1.

information presented, but also to capture longer sequences of visual exploration. The analysis of such captured data presents manifold opportunities to enhance the analysis process. For example, similar to previous work on navigation recommendations for exploring hierarchical graphs [GST13], approaches could be developed to suggest what new parts of the tree should be explored next by the user.

For now, our user focus model considers all elements of the tree (inner nodes and scatter plots). Given sufficient tracking resolution, we may apply the degree-of-interest concept also *locally* within a focused scatter plot. There are numerous ways to heuristically compute interest measures from eye-gaze fixations, fixation sequences, and gaze recurrences. Examples include learning relevance of local patterns, or deducing data groups of interest to a given user or analysis session. In the future, we hope to leverage such information and detect important aspects of the analysis problem at hand (e.g., whether there is exploratory or confirmatory analysis going on in a given session), or detect the level of user expertise. Depending on this infor-

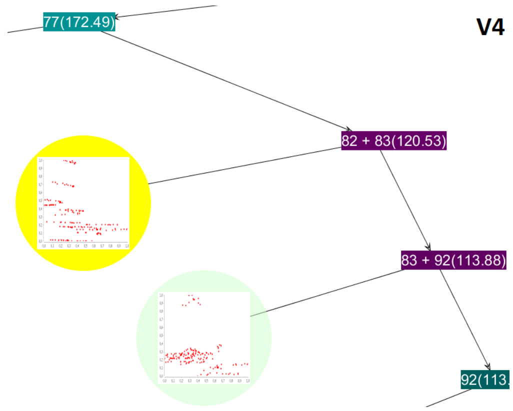


Figure 9: View 4: In a more far location (from V1, V2 and V3) the user notices that there is another interesting scatter plot worth investigation.

mation, the system may adapt its presentation and functionality accordingly. Also, a view recommender module may prevent the user from repetitively going to already examined areas in the display, suggesting instead, previously unseen parts of the view space, based on analysis of the eye-movement (scan) path. To that end, it is interesting to ask how one can do suggestions of other scatter plots to be explored, e.g., based on visual or data-driven similarities.

Another idea is to choose adaptively different visual representations on the nodes based on advanced interest measure computation. For example, in our current visualization approach (scatter plots) we might want to change dynamically between scatter plots, table representation, parallel coordinates visualization, or regression or clusters models computed for a given scatter plot. Also, an interest function should be adaptive also regarding time, e.g., during a longer or repeated analysis cycle. Such interest functions might take into account different objectives, e.g., the user might want to explore the most dissimilar clusters of scatter plots, or explore all scatter plots that have a certain shape. Depending on these objectives the interest function might need to be adjusted.

We also mention that our gaze path visualization could be enhanced to work as an overview tool to represent explored/unexplored regions. A gaze path might eventually serve as a visual history of a whole exploration process. Therefore, we may extend a given gaze path by annotating certain views visited (e.g., scatter plot thumbnails) at certain points in the gaze path (e.g., exceptionally long or short fixation times). Appropriate visual design might communicate a whole analysis session in a single image, which would be a valuable tool for reproducibility and communication of analysis sessions.

6 Conclusion

We presented a concept for visual exploration of hierarchically organized data, that relies on eye-tracking to steer the level of resolution shown. We assume that long gaze fixation times indicates user interest and hence can be used as a proxy to control the visual display of large data. Our effort extends previous work by a new user interface, allowing the navigation and the setting of the degree-of-interest to be determined by eye-movements, and it can be applied on both desktop screens or larger displays (e.g., using wearable eye-trackers). We applied this idea to the specific problem of comparing scatter plot diagrams, and hence support a type of meta-visualization: the elements in the tree are complex objects (visualizations). To this end we applied dendrogram computation based on image features, an approach which may help to overview large amounts of data views by grouping these for similarity. We have shown illustrative use cases for how eye-tracking can enhance a hierarchical data visualization, by mapping eye-gazes to degree-of-interest representations. Yet, our work is in an early stage and we see ample areas for future work. Future work includes high gaze-tracking precision on each node, refinement of interaction operations, view recommending, adaptive visual representations, and analysis provenance visualization. Finally, evaluation of our approach should be done in comparison to non-eye-tracking controls to qualitatively or quantitatively assess strengths and weaknesses of the approach.

References

[ABM07] Wolfgang Aigner, Alessio Bertone, and Silvia Miksch. *Tutorial: Introduction to Visual Analytics*, pages 453–456. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

[Beh16] Lars Behnke. Implementation of an agglomerative hierarchical clustering algorithm in java., 09 2016.

[BKR+14] T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl. State-of-the-Art of Visualization for Eye Tracking Data. In R. Borgo, R. Maciejewski, and I. Viola, editors, *EuroVis - STARs*. The Eurographics Association, 2014.

[CAS01] Mon Chu Chen, John R. Anderson, and Myeong Ho Sohn. What can a mouse cursor tell us more?: Correlation of eye/mouse movements on web browsing. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '01, pages 281–282, New York, NY, USA, 2001. ACM.

[CdART12] Stéphan Clémençon, Héctor de Arazoza, Fabrice Rossi, and Viet-Chi Tran. Hierarchical clustering for graph visualization. *CoRR*, abs/1210.5693, 2012.

[CK15] Daniel Cernea and Andreas Kerren. A survey of technologies on the rise for emotion-enhanced interaction. *Journal of Visual Languages and Computing*, 31(Part A):70–86, 2015.

[Coo06] Lynne Cooke. Is the mouse a" poor man's eye tracker"? *Annual Conference-Society for Technical Communication*, 53:252, 2006.

[EOL14] R. Etemadpour, B. Olk, and L. Linsen. Eye-tracking investigation during visual analysis of projected multidimensional data with 2d scatterplots. In *Information Visualization Theory and Applications (IVAPP), 2014*, pages 233–246, Jan 2014.

[Fur86] G. W. Furnas. Generalized fisheye views. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '86, pages 16–23, New York, USA, 1986. ACM.

[GST13] Stefan Gladisch, Heidrun Schumann, and Christian Tominski. *Navigation Recommendations for Exploring Hierarchical Graphs*, pages 36–47. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[GW03] Joseph H. Goldberg and Anna M. Wichansky. Chapter 23 - eye tracking in usability evaluation: A practitioner's guide. In J. Hyn, R. Radach, and H. Deubel, editors, *The Mind's Eye*, pages 493 – 516. North-Holland, Amsterdam, 2003.

[Hay04] Mary M. Hayhoe. Advances in relating eye movements and cognition. *Infancy*, 6(2):267–274, 2004.

[HMM00] Ivan Herman, Guy Melançon, and M. Scott Marshall. Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 6(1):24–43, January 2000.

[HNA+11] Kenneth Holmqvist, Marcus Nystrom, Richard Andersson, Richard Dewhurst, Halszka Jarodzka, and Joost van de Weijer. *Eye Tracking. A comprehensive guide to methods and measures.* Oxford University Press, 2011.

[JK03] Robert J.K. Jacob and Keith S. Karn. Commentary on section 4 - eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In J. Hyn, R. Radach, and H. Deubel, editors, *The Mind's Eye*, pages 573 – 605. North-Holland, Amsterdam, 2003.

[KMSZ06] Daniel A. Keim, Florian Mansmann, Jorn Schneidewind, and Hartmut Ziegler. Challenges in visual data analysis. In *Proceedings of the Conference on Information Visualization*, IV '06, pages 9–16, Washington, DC, USA, 2006. IEEE Computer Society.

[LR96] John Lamping and Ramana Rao. Visualizing large trees using the hyperbolic browser. In *Conference Companion on Human Factors in Computing Systems*, CHI '96, pages 388–389, New York, NY, USA, 1996. ACM.

[MRS08] Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. *Introduction to Information Retrieval*. Cambridge University Press, New York, NY, USA, 2008.

[OFWB03] J. O'Madadhain, D. Fisher, S. White, and Y. Boey. The JUNG (Java Universal Network/Graph) Framework. Technical report, UCI-ICS, October 2003.

[PB05] Alex Poole and Linden J. Ball. Eye tracking in human-computer interaction and usability research: Current status and future. In *Prospects, Chapter in C. Ghaoui (Ed.): Encyclopedia of Human-Computer Interaction. Pennsylvania: Idea Group, Inc*, 2005.

[PGB02] Catherine Plaisant, Jesse Grosjean, and Benjamin B. Bederson. Spacetree: Supporting exploration in large node link tree, design evolution and empirical evaluation. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis'02)*, pages 57–, Washington, DC, USA, 2002. IEEE Computer Society.

[RFAS08] Kerry Rodden, Xin Fu, Anne Aula, and Ian Spiro. Eye-mouse coordination patterns on web search results pages. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '08, pages 2997–3002, New York, NY, USA, 2008. ACM.

[SSB+15] Lin Shao, Timo Schleicher, Michael Behrisch, Tobias Schreck, Ivan Sipiran, and Daniel A. Keim. Guiding the Exploration of Scatter Plot Data Using Motif-based Interest Measures. *IEEE Int. Symposium on Big Data Visual Analytics*, September 2015.

[SvLS12] Maximilian Scherer, Tatiana von Landesberger, and Tobias Schreck. A Benchmark for Content-Based Retrieval in Bivariate Data Collections. In *Proc. Int. Conference on Theory and Practice of Digital Libraries*, 2012.

[SZCJ16] Popelka Stanislav, Stachon Zdenek, Sasinka Cenek, and Dolezalova Jitka. Eyetribe tracker data accuracy evaluation and its interconnection with hypothesis software for cartographic purposes, 2016.

[vLKS+11] T. von Landesberger, A. Kuijper, T. Schreck, J. Kohlhammer, J. van Wijk, J.-D. Fekete, and D. Fellner. Visual analysis of large graphs: State-of-the-art and future research challenges. *Wiley-Blackwell Computer Graphics Forum*, 30 (6):1719–1749, 2011.

Session 4: Digital Media Experience

HETZI - Jump and Run: Development and Evaluation of a Gesture Controlled Game

Maximilian Schmidt
University of Applied Sciences Ansbach
Ansbach, 91522
maximilian.schmidt@hs-ansbach.de

Cornelius Pöpel
University of Applied Sciences Ansbach
Ansbach, 91522
cornelius.poepel@hs-ansbach.de

Abstract

In this paper an interactive game is described, that has been developed for a company using new technologies for raising awareness at job fairs. An already existing framework of a game terminal for job fairs is used. The corporate identity, the content of the advertised job, but also the attraction of a young and technologically interested audience is combined in a game in order to develop an interactive experience that is easy to grasp, raises awareness for the company and is fun to play. Within the gesture controlled Jump and Run game, the audience accompanies the character "Hetzi" through the guts of a colocation centre, avoiding hazards and animating the player to move in order to win, the company's vision of a modern communication strategy became true. After describing the idea and implementation an evaluation of the game is presented. Due to the results and implemented improvements the company is convinced and the game will be used in future job fairs.

1 Introduction

At job fairs companies generally try to catch the attention of the visitor and want to be remembered by potential future apprentices. In order to achieve this goal for a company providing hosting solutions and being an experienced data centre operator a gesture controlled Jump and Run game was developed. It was implemented on an already existing terminal designed by Fliehr [Fli15], which consists of a Windows Computer running Max 7, a Full HD display in portrait orientation and some additional hardware. The Jump and Run game replaced a version of Tetris in the original terminal application and had to be integrated seamlessly into an existing Max 7 environment, guaranteeing that all other features of the application kept working.

The game was tested and evaluated against four hypotheses in order to be able to figure out weaknesses and strengths of the game and to be able to improve it with respect to the use at job fairs. The hypotheses were defined in a way allowing to figure out whether the game will contribute to the company's overall appearance at fairs. Subsequently final adjustments to further improve the game were made.

Especially in the context of a job fair it is important for a company to be remembered by the visitors. Past experience shows that usually exhibitors do not rely on Jump and Run games for marketing purposes. Therefore it should serve the purpose of becoming a more recognizable brand and helps to make the appearance more memorable. Thus hypothesis one focuses the question whether the player will associate the game with the company in case the game occurs within the company's working environment and the player becomes aware of that.

Hypothesis two: If the sound matches the events on the screen, the sound might contribute to the overall gaming experience and enjoyment.

Hypothesis three: If the controls of the game are simple and intuitive or at least easy to describe and comprehend, the player is able to play the game instantly and without time consuming explanations. This would contribute to making a fast impression on visitors, who usually only spend little time at each single stand of an exhibition.

Hypothesis four: If the complexity of the game is increasing in correlation with the amount of time the player is investing, it equally suits both new and more experienced players.

2 Related Work

Fliehr describes how the single components of the underlying terminal are arranged and function. Moreover, he suggests improving the integrated camera based tracking system and a further enhancement of the brick-based game [Fli15]. Since the version of Tetris and the camera based system were removed, there was no need to consider these improvements.

In: W. Aigner, G. Schmiedl, K. Blumenstein, M. Zeppelzauer (eds.): Proceedings of the 9th Forum Media Technology 2016, St. Pölten, Austria, 24-11-2016, published at http://ceur-ws.org

Lubitz and Krause [Lub12] proved that even if input methods for controlling a Jump and Run game character are considered easy by a very small group of persons, a large group of players may consider the controls too demanding. This emphasizes the importance of usability testing and the necessity of a very simple interaction.

Trepte et al. [Tre11] examined the relation of player performance, game-related self-efficacy experience and the enjoyment of the gaming within an equation. Their study reveals that player performance has a significant influence on game enjoyment. This relation has to be balanced accordingly in order to create an enjoyable experience for both beginners and advanced players.

3 Method

The Jump and Run game was developed in close collaboration with the company. Each step of the development process was shown to the company in order to guarantee the game is tailored according to the company’s vision. After the game requirements had been specified, the game design, sound and controls were developed. Finally the game was implemented on the actual terminal and evaluated.

The evaluation practice consisted of a two step process. During and after the usability test, the participant took part in a narrative interview. Afterwards a questionnaire was filled in by the participant, ensuring the systematic survey of the overall impression of the game. .

The questionnaire was comprised of 19 questions, by which the participants were asked to either rate on a scale from one to five, fill in text or tick a checkbox.

Further insights were gained by the live feedback of the participants while playing the game and the statements they made without prior request.

Ten persons participated in the survey. Most of them were employees of the company the game was designed for. All other participants were been given a short outline of the company.

4 Implementation

The usability test was performed with each participant individually, during which the participants were standing roughly one meter in distance from a screen in front of a Leap Motion controller installed at hip height. After a short introduction to the terminal and how the test had to be performed, the start-screen of the game was presented to the participant.

Followed by an initial instruction of how to play the game by the supervisor the participants were asked to play and explore the game by themselves for a period of at most ten minutes. If a user had any questions, during gameplay, they were answered and used to start a short conversation in order to gain further information.

Once the participants finished playing, they anonymously filled in the questionnaire.

5 Development

The following specifications were outlined at the beginning of the development process:

- The game consists of one endless level, dynamically and randomly put together.
- The speed of the game and of the character increases continuously throughout the game to increase the challenge (which is important to maintain motivation [Har10]).
- Different kinds of obstacles appear randomly.
- The game ends as soon as the character collides with an obstacle and can be restarted multiple times.
- The character is controllable via a simple hand gesture recognized by a Leap Motion controller.
- An 8-bit style melody is playing when the game has started.
- The game is running in an existing Max 7 environment of the terminal.

The storyline of the game follows an employee who has to fix a broken server at the other end of a colocation centre. On his way through the colocation centre he needs to avoid colliding with obstacles in his path.

5.1 Design

The overall design is kept very basic in a “flat” design in order to match with the existing design of the other terminal application and to create a modern, suitable and “clean” look. The game is kept two-dimensional, which lowers the barrier for people who only play occasionally as it creates a feeling of control. [Gan16]

After designing all the single visual components of the game, a sprite sheet was generated[1]. This helped to further reduce the needed resources, improve performances and allowed for quick and small adjustments of the design.

5.1.1 Game character

The game character ("Hetzi") shows a somewhat stereotypical employee of the company. "Hetzi" is wearing "nerdy" glasses, has an inconspicuous haircut, is wearing a black shirt with the company's logo, which is represented by the red "H" and carries a toolbox with the utensils he needs in order to repair a broken server. "Hetzi's" appearance wilfully plays with the widespread notion of the portrayal of an IT-technician (although many might disagree with this representation).

Figure 1: Game character - Hetzi

To make him look a little more innovative and futuristic he is surfing on a hoverboard instead of just running through the colocation centre.

5.1.2 Obstacles

There are three different kinds of obstacles inspired by real objects, which can be found in a data centre

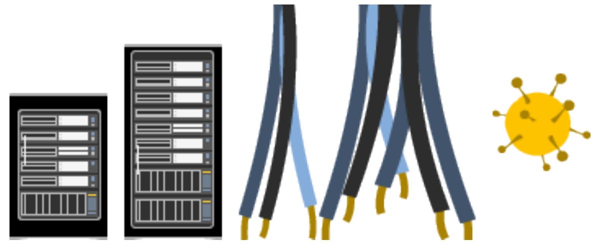


Figure 2: Obstacles

The first types are servers that are modelled in two different sizes. Both servers have the same width but differ in height. The servers are standing on the floor and the character has to avoid a collision with them by jumping.

The second types of obstacles are cables, which (in spite of the usual safety regulations) are hanging from the ceiling and creating a danger for "Hetzi", if he jumps into them.

The third type represents a hacker attack pictured symbolically as a virus. The virus obstacle spawns randomly at one of three predefined heights making it more difficult to avoid a collision with them and creates more diversion.

5.1.3 Setting

The scene of the game shows a colocation centre which is similar to one of the company's (most probably without the hazardous surrounding). One of the special features set in the background is a diagonal ceiling supporting an energy efficient air circulation which is why this feature is represented in the design of the setting. Furthermore there are servers, not too similar to the obstacles, standing in the background in addition with the company's logo on the wall.

Figure 3: Setting

[1] Sprite sheet: An image containing all images used in the game

Obstacles are either standing on the grey floor (servers), hanging from the ceiling (cables) or flying somewhere in between (viruses).

The character has a fixed x-Position (horizontal) and a controllable y-Position (vertical) to allow him to jump. If the character is not jumping he is hovering in a sinusoidal movement above the floor.

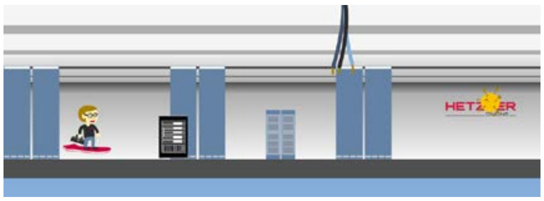


Figure 4: Character and obstacles in the setting

5.1.4 Score

The current score and the high score are displayed close to the top-right corner of the game's canvas. The high score is indicated by "HI" and the current score updates in real-time while playing.

Figure 5: Score and high score

5.2 Sound

Three events result in a sound aiming to create a more interactive gaming experience:

- The start of a jump of the character.
- Reaching a multiple of 100 points.
- A collision of the character and an obstacle.

Moreover a background soundtrack starts every time the game begins. The soundtrack is an 8-bit sound file, which is supposed to be reminiscent of the sound design of older 8-bit-platform games such as Super Mario Bros.®.

5.3 Controls

In order to keep the control as easy as possible only one action can be triggered by the player: The jump. This jump of the character is triggered by a swipe-up with either the right or left hand of the player within the interaction area of the Leap Motion controller.

If the Leap Motion recognizes a swipe-up the y-axis variable increases. If a certain threshold is reached the jump starts. The threshold guarantees that the jump isn't triggered by accident, allowing the player to slightly move their hand without affecting the game. Since the learning curve appears to be very gentle, the game offers entertainment for all possible user groups.

5.4 Software development

The software development process started with the analysis and the adjustment of the pre-existing terminal application. The new game logic was then implemented in JavaScript and integrated in the Max 7 environment, followed by the integration of the Leap Motion Controller into Max 7.

5.4.1 Initial situation

The initial application consisted of:

- The start screen,
- A menu to choose one of two games,
- The Tetris-clone,
- Numberquiz (a game where you have to convert decimal into binary numbers and plug the number with network-cables in a switch with the upper row representing ones and the lower row zeros as fast as possible),
- A photograph of the player, taken by an attached webcam and printed together with the time from the Numberquiz

The navigation of the application was achieved by holding physical Tetris-bricks with a fiducial mark into the field of view of the installed camera. The camera then tracked the fiducial mark and translated the information into the desired output.

5.4.2 Programming Environment Max 7

The Jump and Run game was added to the pre-existing software whilst all other features had to remain working. The navigation links were adjusted accordingly.

To navigate through the application the Leap Motion was introduced. As there is a maximum of two navigational options on each screen one option is triggered by clenching a fist with the left hand for a short amount of time while the other option is triggered by holding a fist with the right hand.

The layout of all the screens has been adjusted to take advantage of the full resolution of the terminal's screen (1080*1920px).

5.4.3 JavaScript in Max 7

For a Jump and Run game the "JSUI-object" fits best which is the equivalent within the Max framework to the Canvas-element in HTML (Cyc16). The JavaScript code updates the position of the background, the character and the obstacles every frame. Then all the elements on the canvas are cleared and redrawn with their new positions.

The character has a fixed x-Position creating the illusion of a forward movement by decreasing the x-Position of all other elements. The hovering of the character is created by updating its y-Position. It increases until it reaches an upper threshold and then decreases until it reaches the lower threshold. When a jump is triggered the y-Position is updated in regard to the existing thresholds.

Initially six random obstacles are created. If an obstacle's x-Position is lower than zero minus the obstacle's width, it is removed and a new obstacle is created. The x-Position becomes then the x-Position of the prior obstacle plus its width and an additional random gap. The virus obstacle can only be created if the speed, which increases throughout the game, has exceeded a certain limit, since it is the most difficult obstacle to avoid a collision with.

The collision detection validates if:

- the character's x-position is smaller than the obstacle's x-position plus its width
  and
- the character's x-position plus its width is greater than the obstacle's x-position
  and
- the character's y-position is smaller than the obstacle's y-position plus its height
  and
- the character's y-position plus its height is greater than the obstacle's y-position

If all conditions are met, it returns a collision and the game state is set to game over.

The background consists of two parts. If the first part's x-position is smaller than zero minus its width, it receives a new x-position which is the second background's x-position plus its width.

The current score is constantly increasing. If the current score modulo 100 is zero the achievement sound is played. If the game state is set to game over and the current score is higher than the high score, the high score gets updated.

If the game state is set to game over, the game frames and the update cycles stop. If the game is then restarted everything besides the high score is set to the initial values and the game begins again.

If you are using "this." in JavaScript as you usually would, problems can occur as Max sometimes falsely seems to relate "this." to the JSUI-object in your patch. This can lead to errors in the code, shown by the console or to unexpected behaviours of the game. Therefore you should avoid using "this." or you would have to double-check each occurrence of "this." to make sure it is working the way you need it to.

5.4.4 Leap Motion integration in Max 7

Using the "leapmotion-object" by Jules Francoise [Fran14] the integration of the Leap Motion Controller is straightforward. The object extracts all the tracking data from the Leap Motion and utilizes it in the relevant Max-Patch.

To navigate through the application the grab strength of both hands is evaluated. If it is found to be higher than a threshold, a variable is increased; if it is lower the value of the variable is set to zero. The navigational option is triggered if the variable reaches a threshold. Additionally, the confidence (a probability value of the placement of an object) of each hand must be evaluated, since the "leapmotion-object" keeps the last known value of the grab strength stored and does not reset it when the tracked object is moved out of the interaction area. While the user moves his hand to the edges of the interaction area, the confidence decreases allowing the program to detect the possible leave of the users hand and therefore is able to set the value of the relevant variable to zero.

To provide a visual feedback to the user a loading bar beneath the navigational option continues to fill while the user is holding a fist. By that way the user is able to experience how long to hold the fist and when to open the hand again.

6 Evaluation

The results of the evaluation were split into four parts:

- Design.
- Gameplay.
- Overall.
- Improvements.

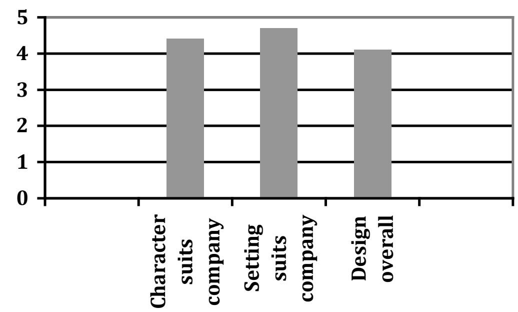
The participants surveyed helped in identifying major bugs as well as receiving an overall first impression from the users. The study was not designed with the goal to generate statistically exact measurement results on an overall user experience but to see tendencies in the user experience with respect to the hypotheses and the aims of the development. Thus the graphs in the following represent the results according to the answers of the participants and are used here as a means of users tendencies in the perception of the game.

Eight out of the ten participants were male. The average age throughout the group was 24.4 years. Two of them had prior experience with a Leap Motion controller. Only one person had never played a Jump and Run game before. At the time of the survey, eight of the ten participants were working at the company the game was designed for.

6.1 Design

Being asked what the character represented, 90% recognized the character as a stereotypical employee of the company. 60% identified his means of transportation as a hoverboard. Even though 40% did not, only one participant wondered why the character was hovering over the floor and thought it to be a "carpet-like thing" out of Aladdin's tales. All participants identified the background as a colocation centre. With all above mentioned characteristics the participants were asked to fill in the blank text. Rating on a scale from 0 (not at all) to 5 (fully agreed) the participants agreed that the character as well as the setting suited the company's conception and image.

Table 1: Design



The overall design was rated 4.1 out of 5 points.

Some of the participants were not sure if the servers in the background of the game were obstacles. At first they had trouble distinguishing the obstacle servers from the background servers. However, after a short period of playing the game this problem disappeared. Nevertheless, the background was adjusted in order to address this problem.

6.2 Gameplay

6.2.1 Sound

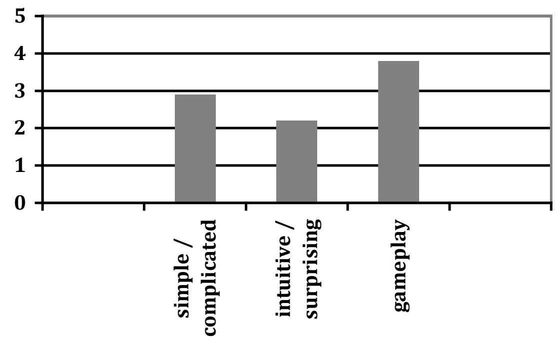
The sound was perceived as either supporting or disturbing. The participants, who perceived the sound as disturbing, agreed on the fact, that the game should contain sound, but they were displeased with the chosen song.

All participants agreed that the sounds for jumping and a collision were fitting and enhanced the quality of feedback.

6.2.2 Controls

On a scale where 3 represented "just right" the swipe up was considered neither as very simple nor as complicated. Most participants said it was rather intuitive and fast to learn.

Table 2: Controls



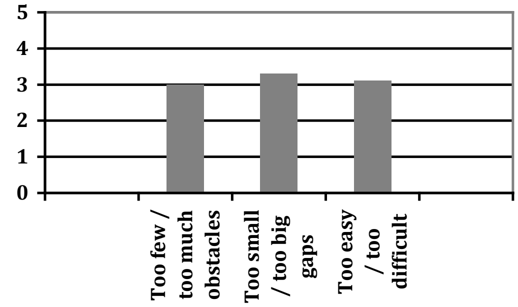
At first the results were quite surprising, because a swipe up did not seem to be challenging at all. The problem most participants experienced resulted from the narrow tolerance range of the Leap Motion controller, which often failed to recognize the movement as a swipe up. Several reasons could be identified. In case the movement was carried out too slow, the participants were able to quickly adapt the movement. However, in case the participants hand had left the interaction area, the participants were not able to realise this and therefore could not adjust their motions accordingly. For that reason further adjustments had to be made to the game (6.4 Improvements).

Despite the difficulties, the participants rated the gameplay 3.8 out of 5 points. This result was used to slightly adjust the game-controls in order to strengthen the gameplay.

6.2.3 Obstacles

Regarding the obstacles, the results of the survey showed, that the participants were content with the amount of obstacles in the game and the positioning of these obstacles as well as the level of challenge for beginners and experienced players. They were asked to rate on a scale from 0 (too few / small / easy) to 3 (just right) to 5 (too many / big / difficult).

Table 3: Obstacles

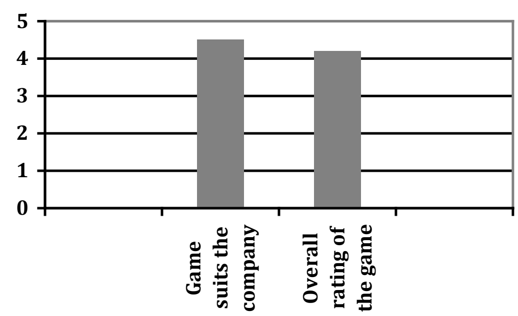


We found that it is rewarding to conduct small adjustments to the individual proportions of the visual elements. The ratio between the character’s height, the possible jump height, the obstacle’s width and height as well as the period of appearance of the obstacles and the speed of the game have to be treated carefully.

6.3 Overall

The participants of the study expressed that the game suited the company very well. Despite some weaknesses they rated the overall game 4.2 out of 5 points.

Table 4: Overall



Considering the difficulties some of the users experienced, this appeared to be a surprisingly high rating, which indicates a possible potential of the game.

6.4 Improvements

After the evaluation, the following improvements were made. The difference between foreground and background was emphasized by slightly desaturating and blurring the background.

The threshold for triggering a jump was lowered. Therefore a jump is triggered faster and easier.

A small green box was included to indicate if the left, right or both hands are within the interaction area of the Leap Motion controller. This should serve as a permanent feedback for the user and makes it easier to avoid mistakes.

Not yet implemented but possible improvements:

- Building a physical boundary for the interaction area of the Leap Motion.
- Replacing the 8-bit song with a more popular melody.

For further adjustments the game would need to be retested by a broader audience at an exhibition where the terminal shall be displayed.

7 Conclusion

Hypothesis one: The participants did associate the game with the company and responded that it was a “very good” suit for the company. We believe that this is mostly the result of the design of the character, the obstacles and the environment the game plays in.

Hypothesis two: For the participants of the study sound played an important role in making the game enjoyable. Even if the sound does not match with the player’s idea of a suitable music, participants indicated that it contributes to the overall gaming experience and enjoyment.

Hypothesis three: If the controls of the game are simple and intuitive, participants were able to instantly play the game. Two of the participants came back after the usability test and still remembered the gesture. We conclude that once an intuitive gesture is performed a few times, it will not be forgotten quickly and the game can be enjoyed more easily.

Hypothesis four: As the game becomes more difficult the longer it runs, newbies as well as more experienced participants enjoyed the game. The only difference seemed to be that more experienced participants made faster progress in the game and reached a higher speed in fewer attempts.

8 Prospect

As mentioned in section 6.4. a retest within the framework of an exhibition or ideally a job fair including a bigger number of participants of the target group is a future project to be undertaken.

According to our experience the most important part of making a gesture-controlled Jump and Run game enjoyable is to give the user constant feedback about the precision of the applied gesture. This should be done through the interface the user sees on the screen rather than a person who constantly has to correct the player’s movement.

Especially building a physical boundary to indicate the interaction area of the Leap Motion controller would be a great benefit for any gesture controlled application.

As the Jump and Run game is mostly written in JavaScript it could be adapted to be used on the company’s website or implemented as a smartphone app – which three participants of the usability test also suggested.

Developing a lightweight casual game with easy controls can be a great benefit for a company. It increases the attractiveness of an exhibition stand and can be used as a giveaway in the form of a smartphone app enhancing a company’s impact on future applicants and raising memorability.

A possible future improvement on the overall experience would be the implementation of a persistent high score list.

Furthermore it would be interesting to implement different input methods for the jump like a Wii balance board or a Kinect camera recognising a jumping user or just a simple enter key.

References

[Cyc16] Cycling '74. *Cycling '74 Forum*. 2016. 28. 07 2016. <https://cycling74.com/forums/topic/javascript-performance-vs-max-objects/#.V5qTb7iLS00>.

[Fli15] Fliehr, Andreas; Müller, Anna; Ehrenfeld, Franziska; Pöpel, Cornelius. „Haptische Interaktion beim Messeauftritt: Ein audiovisuelles Gameterminal zur spielebasierten Attraktivitätssteigerung einer Firmenpräsentation.“ Editor G. Schmiedl. Glückstadt: Verlag Hülsbusch, 2015.

[Fran14] Francoise, Jules. *Leap Motion skeletal tracking in Max*. November 2014. Leap Motion skeletal tracking in Max. 28. 07 2016. <http://julesfrancoise.com/leapmotion/>.

[Gan16] Gandolfi, Enrico. „The two dimensions as a metaphor of control in gaming landscapes." 2013. *Game - the Italian journal of game studies.* 28. 07 2016. <http://www.gamejournal.it/the-two-dimensions-as-a-metaphor-of-control-in-gaming-landscapes/>.

[Har10] Harrigan, Kevin A.Dixon, Michael J.; Fugelsang, Jonathan A. „Addictive Gameplay: What Casual Game Designers CanLearn from Slot Machine Research." *Vancouver Digital Week*. Vancouver: ACM, 2010.

[Lub12] Lubitz, Kolja und Markus Krause. „Exploring User Input Metaphors for Jump and Run Games on Mobile Devices." Springer Berlin Heidelberg, 2012. 473-475.

[Tre11] Trepte, Sabine; Reinecke, Leonard. „The Pleasures of Success: Game-Related Efficacy." *CYBERPSYCHOLOGY, BEHAVIOR, AND SOCIAL NETWORKING*. Edition. 14. 9. 2011.

Embodied Material Guidance: Augmenting Material for Carving

Marcel Penz
Umeå Institute of Design
Umeå University
SE-90187 Umeå, Sweden
marcel_penz@hotmail.com

Sreyan Ghosh
National Institute of Design, India
R & D Campus,
Bangalore 560 022, Karnataka, India
contact@sreyan.com

Abstract

When learning a new skill, video tutorials are often the first choice to get help. For creative domains like carving, this is often not sufficient as the desired expression require reacting on changes in the form constantly. A crucial part of the making process is a close observation of the material. Looking at various domains we have chosen wood carving, to propose new forms of interactions that can support creative learning processes. Instead of focusing on tools we aim to augment the material for the novice carver. The concept aims to reduce the fear of learning by replacing traditional carving material (like wood) with 3D printed material. Guides on where to carve are part of the printed base material and form embodied guidance. We assume artistic self expression as an additional goal for carving and provide added levels of randomness to the material structure through variable colors and material densities while using Virtual Reality (VR) to provide a change in perspective.

1 Introduction

The success of numerous online tutorials for making indicates how people gather first hand insights in learning a new skill. Free online platforms for learning such as edx.org, coursera.org or others offer a variety of courses for both hands-on and theoretical knowledge building conveyed through video. Often tutorials allow for copying a stepwise process to achieve a certain goal. For artistic activities like painting, clay modelling or wood carving static videos – which don't adapt to the learner's context – are only of limited help. For painting different interactive systems have been created that have the potential to support the learner by enhancing brush and/or canvas by digital components (Otsuki et al., 2010, Vandoren et al., 2009). Also, for crafting three dimensional models, hybrid approaches that combine digital and physical components are being developed (Golsteijn et al., 2013, Oe, Shizuki & Tanaka, 2012). However augmenting or enhancing the craft material itself is not very well explored. Often the main focus is to support building digital models by interacting with physical material (Reed, 2009, Smith, Thomas & Piekarski, 2008). In our work we focus on the craft's material itself and how a user can be supported in shaping the material. The context is carving where we assume that 3D printing material suited for carving can be created. In a different context researchers are already working on wood-inspired 3D printed material (Compton & Lewis, 2014).

2 Research & Ideation

2.1 Ideation Workshop

In order to better scope the project we invited design students to participate in an ideation workshop revolving around the question of how the future home can be a place for collaborative making and learning through making. During the 45 min workshop, six students of the Master's programme in Interaction Design at the Umeå Institute of Design were asked to follow a set of tasks involving reflecting on hobbies and their tools and how the tools will change in the future. Some participants found it difficult to force features onto tools that should enable creativity. During the final discussion a common point of view was created stressing that tools should be enablers rather than means for creativity. By not focussing on the tools we were inspired to think about the material instead. Hence we decided to take an approach of augmenting the medium i.e, the carving material.

2.2 Research on Tutorial Videos for Making

To identify the issues that the hobbyist carver faces, we reviewed tutorial videos on the web (Farnsworth, 2014, Schrettl, 2014) trying to place ourselves in the mindset of a beginner in that field. Our reflections towards these videos helped us identify a few issues. When staying in the virtual world, support through video

In: W. Aigner, G. Schmiedl, K. Blumenstein, M. Zeppelzauer (eds.): Proceedings of the 9th Forum Media Technology 2016, St. Pölten, Austria, 24-11-2016, published at http://ceur-ws.org

seems to work effectively e.g. for digital tools such as creative software (Photoshop, 3DS Max). We assumed that the flat representations on the screen perform inadequately when it came to learning with actual tangible material. Videos are sufficient for defined tasks or step-by-step instructions. However, videos do not work as well for artistic expression such as carving due to the fact that the material in the user's hand will always be perceived differently than the one displayed. The biggest challenge for the novice carvers seemed to be in the ability to understand when to stop carving.

3 Concept

"Every block of stone has a statue inside it and it is the task of the sculptor to discover it" — *Michelangelo Buonarroti.*

In the same way a blank page can be a barrier for a writer a block of wood can be an issue for a sculptor. Once this issue is overcome, the process of turning this block of wood into a small figurine without using machinery requires both craft skills and skills for creative expression. In our concept we propose using 3D printing to create a specialized base material for carving embodied guidance. In this way we want to guide the carver towards a shape that is already part of the carving block and foster a dialog with the carving material through this augmentation. In addition we use VR as part of supporting creative reflection.

3.1 Scenario

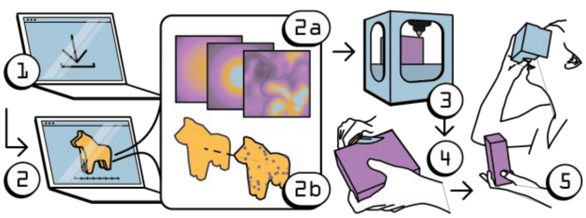


Figure 1: Diagram showing the different steps of the concept from downloading a 3D model to carving it out and using VR to change perspective

The user is inspired to carve out the iconic Swedish Dala horse (Swedish: Dalahäst) and downloads or builds her own CAD model (see fig. 1.1). In a software similar to current 3D printing software, she imports the model that is embedded in a virtual cube (see fig. 1.2). Control elements let her choose how much of the horse should be visible when later cutting through the material. This is defined through 1.), contrast in color between the horse and the waste material and 2.) through guiding dots aligned with the edge of the horse (see fig. 1.2a, 1.2b and fig. 2). The user should also be able to add additional color gradients (see fig. 1.2a). A sensor unit is embedded in the print, able to communicate with a VR headset. After the block is 3D printed (see fig. 1.3) she can carve the print immediately (see fig. 1.4). The more she cuts, the more guiding dots appear making her aware where to carve away more material. As the color of the material changes while cutting she is inspired to modify the final shape for aesthetic reasons. To achieve a better understanding of the proportions of the object inside, she can use the sensor unit in her workpiece to connect to VR headset using wireless technology. By switching to VR she gets to see her work from a different perspective (see fig. 1.5).

3.2 Embodied Guidance: Guiding Dots, Color Gradients and various Levels of Material Density

Figure 2: Left: Cut out of the 3D printed material, middle: guiding dots at the surface of the horse, right: color gradients to indicate proximity to the final shape

When creating a drawing people use techniques like tracing paper as support. The supporting techniques for a novice carver creating 3D shapes are limited. Often carvers draw guiding lines on the workpiece that are carved away immediately, hence using 2D solutions for a 3D problem. When interviewing makers we discovered that – especially for beginners – there is often the fear of carving away too much from the carving material. By printing guiding dots in the base material, drawing lines becomes obsolete and confidence is fostered. As the novice carver starts to carve out the printed block, a visual imprint of the horse inside the block emerges. This acts as feedback that helps the carver to stop carving in time, thus overcoming the challenge of carving in too deep. Further for intermediate users we envision various levels of support by limiting the guides indicating the border of the shape adjusted in software (see fig. 2, center image). In this way we aim to support the learning process. We also envisioned embedding a color gradient changing the color from a surrounding boundary box to the object inside. When the novice carver sees the change in color, she is made aware that the enclosed object is starting to surface (see fig. 2, right image). Other ways of showing when the carver gets close to the object could be to use variable densities of material, a property that can be set for 3D printers. By increasing the density of the object inside (the Dala Horse) as compared to the surrounding carving block tactile feedback can be provided.

3.3 Creative Reflection on the Material

We focus on the relationship between the user and the material, almost trying to allow for a reflective conversation, where the material provides rich feedback on how it is treated. To achieve this, we propose change in material through adding randomness in the color gradients and providing the option to change perspective in VR.

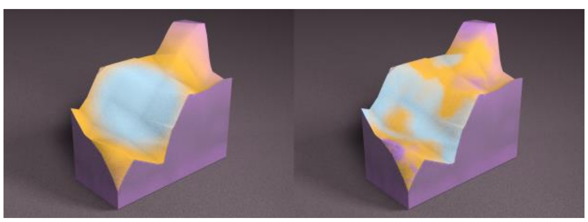


Figure 3: Different turbulences within the gradient

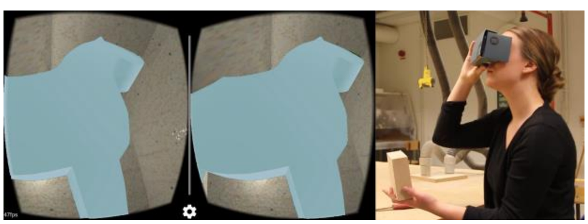


Figure 4: View from the inside of the carving block

3.3.1 Fostering Creativity by adding Randomness

As described before guides can be printed in form of guiding dots or color gradients. However the guiding color gradient doesn't have to be linear. By enabling the user to add gradient turbulences, the embedded object becomes slightly distorted. A low level of color turbulences will lead to uniform coloring. Higher levels of turbulences will cause multi-color 3D gradient patterns (see fig. 3). In the same way material density could change as well. This is an attempt to mimic natural material qualities characteristic to wood. The same way the wood's growth direction can inspire the wood carver to carve different shapes, artificial randomness can hint towards a shape that the creator has not thought of before.

3.3.2 Changing Perspective: Using Virtual Reality

In a last step we experimented with the virtual representation of the initial shape using VR. A good practice in (abstract) painting is to physically step back from the painting to see how the latest brush strokes have changed the overall composition. It is a state of observation and reflection that allows for constant acting and reacting on the painting. Other than this example of a two dimensional canvas, the 3D object allows us to explore the material also from the inside.

With the help of a virtual reality headset the user can place herself inside the box with the object in front of her. This makes her aware of the parts that are not carved out yet and may open up for changing the direction she wants to pursue. By adding a small inertial measurement unit (IMU) the workpieces orientation can be tracked and transmitted to a VR headset using wireless technology. The physical model acts as controller for the user to navigate in 3D space, where the user can move around the virtual environment by moving the physical object (see fig. 4). We created a prototype in Unity & Arduino to communicate ways to navigate in virtual space and presented it to people (see next section). Since the user holds the partially carved out object in her hand she gets rich tactile feedback while at the same time seeing the model from the inside. This disruptive experience might encourage the user to change the way she thinks about the workpiece. After establishing the experience of the inside view, the user will return to the physical world continuing to carve, having the possibility to revisit the virtual space at any time during the process.

4 Evaluation & Reflection

4.1 Interviews

We proceeded to interview four makers to gain qualitative feedback on our concept. The interview participants were chosen based on their background and skill in craftsmanship. We interviewed a teacher for making physical prototypes, a product design student and passionate maker, a prototyper in a design consultancy who likes to tinker in his free-time, and an owner of a 3D printing shop. The participants were asked questions around their personal experience with making at the beginning of each interview. Some recurring topics in all the conversations were:

Sources of inspiration and the reason for making including: Solving a specific problem through making, inspiration by different materials or inspiration through finding interesting projects (online). We also learned that the **tool plays different roles for different people** ranging from being the center of the activity to being mere means to achieve a goal. Lastly we got to know different drivers for making that can either be **outcome driven vs. process driven** as well as making as purposeful vs. making as artistic expression. Especially the distinction between drivers has implications for our design, as our concept is not meant to focus on creating to fulfill a need but is a learning tool for expression in 3D physical objects. When presented with the concept the interviewees saw great value in the material having an embodied object and agreed that the concept could definitely work to train and encourage people to carve. The concept of variable gradients as a guide to stop carving and start sanding was highlighted positively. While the prototyper and the product designer saw value in using VR to gain a different perspective, the teacher did not agree and stated that the shift in medium causes disconnect and the user should not be made to leave the material at any stage of the carving process. The print shop owner suggested that the density of material should also vary along a gradient along with the colors.

4.2 Technical Limitations

Adding the virtual reality feature would lead to multiple problems. The first one concerns the sensor unit embedded in the work piece. To our knowledge it is currently not possible to print the sensor at the same time as the carving block. One way to work around this could be to add the electronics manually in the printing process. The second problem is the changing shape of the module that cannot be tracked easily with available sensor technology. However research in this direction has already been conducted (Reed, 2009, Smith, Thomas & Piekarski, 2008).

To our knowledge there is currently no material available that supports carving. However, when interviewing an 3D printing expert he felt confident that this material could be manufactured easily. Also research suggests that this kind of material can be produced (Compton & Lewis, 2014).

4.3 Beyond Carving

We feel confident that there are advantages in augmenting material by adding additional information as part of the manufacturing process. Changes within the material can provide extra information e.g. considering abrasion and wear of material or inside predetermined breaking points. We hope that this concept inspires others to explore new possibilities on how to physically augment material.

5 Conclusion

In this paper we present a concept to support novice to intermediate skilled people interested in carving. As a main incentive to carve we assume artistic self expression and the learning/ skill building of carving as handicraft. Our proposal focuses on the dialog between the learner and the material and how the material can be augmented to be more expressive. We show this in three parts of our concept: First, the materialization of guidance manifested in a 3D printed block (embodied guidance). This allows the user to fully concentrate on the material and the learning process. The embodied guides aim to reduce fear of making mistakes and to encourage training, since the final product is already part of the workpiece. Second, to support for more creative exploration of the material we envision the functionality to change the inner structure (color gradients and density) of the material by adding randomness. Both embodied guidance and the creative reflection with the material allow learning how to carve by making use of the digitally enhanced material while staying in the physical world interacting with it in a tangible way. Third, we offer a way to use the digital realm for changing the perspective to see inside the physical workpiece using VR. We created prototypes used in discussions with four makers. In those discussions we found that the makers were generally enthusiastic regarding various levels of support and they could see value in random changes of the material structure. However, not everyone agreed on the importance of switching between VR and physical world while carving.

References

Compton, B. G. and Lewis, J. A. (2014), 3D-Printing of Lightweight Cellular Composites. Adv. Mater., 26: 5930–5935. doi:10.1002/adma.201401804

Farnsworth, J. (2014). Wood Carving Tools & Techniques for Beginners. Youtube, Published on Jul 10, 2014. Last-accessed 25. Aug, 2016

Golsteijn, C., Hoven, E. V., Frohlich, D., & Sellen, A. (2013). Hybrid crafting: Towards an integrated practice of crafting with physical and digital components. Pers Ubiquit Comput Personal and Ubiquitous Computing, 18(3), 593-611. doi:10.1007/s00779-013-0684-9

Oe, T., Shizuki, B., & Tanaka, J. (2012). Scan modeling. Proceedings of the 10th Asia Pacific Conference on Computer Human Interaction - APCHI '12. doi:10.1145/2350046.2350096

Otsuki, M., Sugihara, K., Kimura, A., Shibata, F., & Tamura, H. (2010). MAI painting brush. Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology - UIST '10. doi:10.1145/1866029.1866045

Reed, M. (2009). Prototyping digital clay as an active material. Proceedings of the 3rd International Conference on Tangible and Embedded Interaction - TEI '09. doi:10.1145/1517664.1517733

Schrettl, P. (2014). Schnitzen 1 kurz. Youtube. Published on Aug 30, 2014. Last-accessed 25. Aug, 2016

Smith, R. T., Thomas, B. H., & Piekarski, W. (2008). Digital foam interaction techniques for 3D modeling. Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology - VRST '08. doi:10.1145/1450579.1450592

Vandoren, P., Claesen, L., Laerhoven, T. V., Taelman, J., Raymaekers, C., Flerackers, E., & Reeth, F. V. (2009). FluidPaint. Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces - ITS '09. doi:10.1145/1731903.1731914

All Around Audio Symposium

SAMBA – Smart Data for Music Business Administration
Andreas Gebesmair, St. Pölten UAS, Austria

Challenges and Problems of Automated Mastering
Patrik Lechner, St. Pölten UAS, Austria

Interactive Storytelling & Audio Apps for HbbTV
Nicolas Patz, Rundfunk Berlin Brandenburg, Germany

Combining Sonification and Visualization for the Analysis of Business Process Data
Tobias Hildebrand, University of Vienna, Austria

Sound Quality and User Experience
Cornelius Pöpel, Ansbach UAS, Germany

Music Content Recognition with Machine Learning in Academia and Startups
Thomas Lidy, TU Wien, Musimap, Austria
Audio in Digital Healthcare
Jakob Doppler, St. Pölten UAS, Austria

Music Computing from Opera to Avant-garde
Oliver Hödl, University of Vienna, Austria

Electronic Music Archeology
Reinhold Friedl, zeitkratzer, Goldsmith University London, UK

Fluid Control – Media Evolution in Water
Christoph Theiler, wechselstrom, Vienna, Austria

Sounddesign Education
Jörg Lensing, Dortmund UAS, Germany

Speaking about Sound: A Brief Overview of Existing Research on Verbal Timbre Description
Peter Plessas, IEM Graz, MDW Vienna, St. Poelten UAS, Austria

Enabling Realtime Collaboration for Musicians
Helmut Herglotz, sofasession, Vienna, Austria
Variable Acoustics and 3D Audio with Amadeus Active Acoustics
Fabio Kaiser, Rohde BeSB, Salzburg, Austria

Non-conventional Speaker Settings and Multi-channel Architectures
Paul Modler, Hochschule für Gestaltung Karlsruhe, Germany

ParametricSpeakers – Personal Experiments and Possible Markets
Mauricio Valdés San Emeterio, guest artist @ Institute for media archeology St. Poelten