

WebSci'18

Interdisciplinary PhD Symposium



27 May 2018, VU Amsterdam, The Netherlands

Chairs & Editors:

Pete Burnap, Cardiff University, UK

Leslie Carr, Web Science Institute, University of Southampton, UK

Mark Weal, Web Science Institute, University of Southampton, UK

Ivar Vermeulen, Network Institute, VU Amsterdam, The Netherlands

Copyright notice: all rights of the papers, posters, abstracts, etc., contained in the WebSci'18 Events Proceedings rest with their respective authors; the copyright for the events as a whole with the respective chairs/editors.

WebSci'18 Interdisciplinary PhD Symposium

Pete Burnap, Leslie Carr, Mark Weal, Ivar Vermeulen

Sunday 27 May 2018

Tackling the digital divide in UK schools: Young People's perspective on digital literacy education and digital use practices

Charlotte Oliver

Towards a data-driven social machine for online academics' profiles

Bartosz Paszcza

The digital divide in Amsterdam: the perspective and needs of families with a low socio-economic position

Nicole Goedhart

Cyber Hate against immigrants: an Italian case

Komal Florio

An explanation of the changing relations between digital platforms and entrepreneurship in Trinidad and Tobago

Keisha Taylor

Crowdsourcing Application for Responsible Production in Africa

André Baart

Online participation processes in the Green Party Germany

Gefion Thuermer

Tag Thunder: Towards new oral reading strategies for visually impaired people

Judith Jeyafreeda Andrew

A Study on the Emergence of Online Bar Exam Training in China

Xingfang Li

Information Access for Development: Weather Information Access for Rural Ghana

Francis Dittoh

Tackling the digital divide in UK schools: Young People's perspective on digital literacy education and digital use practices.

Charlotte Oliver
University of Huddersfield

Charlotte.Oliver@hud.ac.uk

ABSTRACT

"Today, because of rapid economic and social change, schools have to prepare students for jobs that have not yet been created, technologies that have not yet been invented and problems that we don't yet know will arise."

(Schleicher, 2010, p.42).

The focus of this study is the emerging topic of Digital Literacy that is beginning to be taught in schools across the UK. The definition of Digital Literacy is not yet concrete, as technology develops so to do the skills required to understand, engage and utilise it. Currently, the term can refer to instrumental digital skills (Livingstone et al., 2011) such as how to switch a laptop on or how to use a search engine but the term also encompasses social digital skills (Livingstone et al., 2011), how to communicate effectively online for example, critical digital skills (Polizzi, 2017), including the ability to recognise deceitful information or individuals online, and creative digital (Eshet-Akakai, 2012) skills such as the ability to upload a blog post. Current definitions of what it means to be digitally literate have been informed by academics, professionals and parents in the main and focuses on areas of educational development, safety and capital gaining activity. Although Children may have been consulted on their digital practices there appears to be a current gap in knowledge regarding what digital literacy means to children, for example: how would they define the term? Which digital skills do they think are important? And how would they like digital literacy education in the classroom to be delivered? This research aims to answer these questions by using a collaborative, participatory methodology with year 7 pupils in a UK school, to gain a qualitative insight in to what the children's perspective of the topic is. Supporting the children to articulate their experience of formal and informal digital literacy education and to have their voice contribute meaningfully to knowledge on a topic that directly effects them.

Keywords

Digital literacy, education, participation, children, young people, UK.

1. INTRODUCTION

This study aims to use a Participatory Research methodology, with year 7 pupils aged 11-12 years, to qualitatively explore young people's experience of digital literacy education they have revived thus far, as well as their expectations of what they believe they will receive as they progress through key stage 3. The research will also allow young people to suggest ways they feel digital literacy education delivery in the classroom could be improved, by drawing on their actual experience (including formal and informal) of digital technology. The research will allow for an in depth and nuanced discourse with children about their engagement with digital technologies and participation in online activities, including their social media use. Using a collaborative approach, the project will bring together children and educators to work in partnership, to learn from each other offering a unique and insightful contribution to knowledge.

2. Background

2.1 Digital Divide

Due to their seemingly universal access to digital technology it is often assumed that young people aged 0-18 years fall under the description of the 'Digital Native' (Prensky, 2001, p.2) with the belief that being surrounded by digital devices is enough exposure for them to master the use of such technology, without much or any educational intervention. Over the years Prensky's work has been contested by many scholars (Gui & Argentin, 2011; Helsper, 2017; Lindblom & Räsänen, 2017; Livingstone & Helsper, 2009; Robinson, 2009; Smith et al., 2012; Wei & Hindman, 2011) who have argued digital skills are not universal among younger generations and that there exists a three tiered digital divide (Wei et al., 2011) The first tier separates those who have physical access to digital devices and those who do not, as whilst the majority of UK homes are now online, 10% still do not have internet access in the home, (Office for National Statistics, 2017, p.3) this can be down to socioeconomic status or voluntary abstinence (Helsper, 2017; smith, 2012). The second tier digital divide concerns digital capability and highlights the fact that not all those who have access to digital technology have the same level of skills or knowledge regarding use of that technology (Gui

& Argentin, 2011), in Prensky's work it is assumed access alone is catalyst enough to become an expert or 'digital native' but research has shown these skills still need to be taught or learnt somehow. Current research shows that parents with strong digital skills and peers with digital skills and motivation are the main providers of such education (Meneses & Mominó, 2010; De Almeida, Alves et al., 2011; Festl, 2018), meaning if a child's parents and peers are not digitally confident or motivated it is likely the child's digital skill level will be affected. Similarly, in schools a child's digital skills education is dependent on the school's dedication to providing digital skills training (PSHE Association, 2018), individual teachers skills and motivations (Preston et al., 2014) and can also be effected by the child's motivation to learn such skills (Sims, 2013).

2.2 Digital Literacy

The digital skills associated with the latter two tiers of the digital divide spectrum are often referred to as 'digital literacy skills', this term has no concrete definition but a number of scholars and policy makers have provided definitions they feel fit. Eshet-Alkalai (2004, p.267) suggested that there are five types of literacies that are incorporated within the term 'digital literacy':

(i) photo-visual literacy - learning-to-read from visuals. It is a cognitive skill that uses "vision to think".

(ii) reproduction literacy - the art of creative duplication. This involves the use of digital tools that have the capabilities to edit or combine/recombine new and pre-existing materials (text, audio, video, images) into new works of art or writing.

(iii) branching literacy - the use of hypertext in the creation of non-linear medium of information and the ability to navigate through the displayed information freely. Branching-literate individuals have good spatial orientation and the ability to create mental models, concept maps and other forms of abstract representations in hypermedial environments

(iv) information literacy - the literacy associated with critical thinking and the ability to search, locate and assess web-based information effectively.

(v) socio-emotional literacy - the literacy associated with the emotional and social aspects of online socialising, collaborating and undertaking day-to-day chores e.g. banking and purchasing online. It requires the ability to be highly critical and analytical, to avoid online 'traps', for example being able to identify pretentious people in the chat rooms and avoiding hoaxes and viruses.'

2.3 Education Policy

Whilst over the years individual schools in the UK may have attempted to implement digital literacy education within their teaching either within specialist computer science classes, social science classes, citizenship or PHSE it is not currently a statutory requirement for schools to teach digital literacy beyond the instrumental (Department for Education, 2013) This has prompted a recent report from the House of Lords Communication Committee recommending (2017, p.1)

'that digital literacy sit alongside reading, writing and mathematics as the fourth pillar of a child's education; and that no

child should leave school without a well-rounded understanding of the digital world. Schools should teach online responsibilities, social norms and risks as part of mandatory, Ofsted-inspected Personal, Social, Health and Economic (PSHE) education, designed to look broadly at the issues that children face online.'

Prompting a government green paper, released in October 2017, stating that the government are 'Giving Careful consideration' (HM Government, 2017, p.26). on the topic of requiring all schools to teach PHSE as a compulsory subject including modules on digital engagement.

3. Methodology

The proposed research methodology for this project is a participatory, qualitative approach, that will be co-designed, conducted and analysed by myself as the researcher, up to 30 year 7 pupils and their teacher.

4. REFERENCES

- [1] Gui, M., & Argentin, G. (2011). Digital skills of internet natives: Different forms of digital literacy in a random sample of northern Italian high school students. *New Media & Society*, 13(6), 963-980. <http://dx.doi.org/10.1177/1461444810389751>
- [2] Prensky, M. (2001). Digital Natives, Digital Immigrants Part 1. *On The Horizon*, 9(5), 1-6. <http://dx.doi.org/10.1108/10748120110424816>
- [3] Eshet-Alkalai, Y. (2012). Thinking in the Digital Era: A Revised Model for Digital Literacy. *Issues In Informing Science And Information Technology*, 9, 267-276. <http://dx.doi.org/10.28945/1621>
- [4] Polizzi, G. (2017). Critical digital literacy: ten key readings for our distrustful media age. LSE Media Policy Project Blog. Retrieved from <http://blogs.lse.ac.uk/mediapolicyproject/2017/12/15/critical-digital-literacy-ten-key-readings-for-our-distrustful-media-age/>
- [5] Livingstone, S., Sonck, N., Kuiper, E., & de Haan, J. (2011). Digital literacy and safety skills (p. 2). London: LSE.
- [6] Schleicher, A. (2011). The case for 21st century learning. *The OECD Observer, Suppl. OECD Yearboo*, (282/283.), 42-43
- [7] Helsper. (2016). SLIPPING THROUGH THE NET Are disadvantaged young people being left further behind in the digital era?. London.
- [8] Office for National Statistics. (2017). Internet access – households and individuals (p. 3). London: Office for National Statistics.
- [9] Meneses, J., & Mominó, J. (2010). Putting Digital Literacy in Practice: How Schools Contribute to Digital Inclusion in the Network Society. *The Information Society*, 26(3), 197-208. <http://dx.doi.org/10.1080/01972241003712231>
- [10] HM Government. (2017). *Internet Safety Strategy – Green paper*. London.

Towards a data-driven social machine for online academics' profiles

Submission for PhD Symposium

Bartosz Paszcza

Web Science DTC, Electronics and Computer Science,
University of Southampton
Southampton
bp11g12@soton.ac.uk

Stevan Harnad

Web Science, Electronics and Computer Science,
University of Southampton
Southampton
bp11g12@soton.ac.uk

Les Carr

Web Science Institute, Electronics and Computer Science,
University of Southampton
Southampton
lac@ecs.soton.ac.uk

Jeremy G. Frey

Faculty of Natural and Environmental Sciences, University
of Southampton
Southampton
J.G.Frey@soton.ac.uk

ABSTRACT

The project focuses on availability, scope and quality of scholarly communications data available online and its' potential usage for innovating the way researchers build their online profile. It describes a novel, semi-automatic methodology for evaluation of available data sources. Furthermore, the report describes future work plan aiming to propose a design of a social machine, combining data, computational methods and human input for creation of a complete researchers' online profiles.

KEYWORDS

bibliographic databases, scholarly communication, scientists' online profiles, database analysis, scientometrics

ACM Reference Format:

Bartosz Paszcza, Les Carr, Stevan Harnad, and Jeremy G. Frey. 2018. Towards a data-driven social machine for online academics' profiles: Submission for PhD Symposium. In *Proceedings of ACM Web Science Conference (WEB-SCI'18)*. ACM, New York, NY, USA, Article -, 2 pages. <https://doi.org/>

1 INTRODUCTION

The rapid growth in variety and volume of online scholarly communication data sources happening since the advent of the Web has inspired a body of research and entrepreneurial activities.

In the recent years, we have seen an increase in number of portals distributing such data, and furthermore - a rapid growth in variety of the data collection, processing processes and formats. We have experienced addition of completely new types of data, such as: institutional and disciplinary Open Access repositories (allowing not only to access publication's metadata, but also full text on a large scale) and altmetric portals (supplementing citation statistics with metrics of downloads, mentions in policy documents,

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

WEBSCI'18, May 2018, Amsterdam, The Netherlands

© 2018 Copyright held by the owner/author(s).

media, blogs or social media) [1, 3, 11]. Secondly, contrary the traditional databases using publisher-provided information, we have observed introduction of Web-specific technologies, e.g. usage of web crawlers to find and identify academic documents [4].

Another important trend is the overall disruption of scale happening in scholarly communications. One has to remember that the WWW in its original creation was meant to primarily improve the speed and convenience of circulation of research data and documents [2]. It has indeed deeply transformed scholarly communications: the research now relies on information that can be found on the Web, the exchange of scientific outputs happens primarily online. As an example, the number of research articles published annually has gone from an estimated level of less than 1,000,000 in the end of 1980s to an estimated level of 7,000,000 in 2014 [10]. The "Researchers of tomorrow" report by JISC [6] found out that 30% of PhD students at that time relied on Google and Google Scholar as primary sources of literature search. There is around 100 new search tools (platforms) for researchers that are being created year by year - whereas in early 1990s it used to be less than 10 per year. The availability of data, novel computational methods and fast communication enables a dramatic change in the way researchers communicate and publish their results.

2 THE QUESTION OF RELIABILITY: EVALUATION OF SCHOLARLY COMMUNICATIONS DATA SOURCES

However, the rapid growth of ways of gathering information for research - and research evaluation - creates a necessity to be able to evaluate the scope, quality and characteristics of the sources used. To enable drawing certain inferences from the data, author needs a thorough understanding of the origin of the datasets (sources, methods of data-gathering), data processing, limitations of use, as well as the scope and quality [12].

Hence the popularity of Web data science research needs to be complemented by development of methodology for complex evaluation of data sources, such that will help to understand the opportunities, but also limitations of the datasets used. Lack of

common rules of pre-research analysis of data indicates that the methodology of conducting studies based on the data available from the Web has not kept up the pace. As an example, despite the calls to present basic statistics regarding the methods of finding cross-database overlap, many studies in the area of scientometrics do not contain such information [5]. Furthermore, creation of large number of services creates issues with interoperability, hampering the opportunity for obtaining a more complete picture by combining usage of a number of services to answer research questions.

In my PhD research, I aim to tackle both of the above-mentioned issues. My overarching goal is to attempt to propose a design combining data already available, computational methods and human input for creation of open, interoperable online researchers' profiles.

3 NOVEL METHODOLOGY FOR EVALUATION OF SCHOLARLY COMMUNICATIONS DATA SOURCES

However, in order to do so, I have first taken a look at the methods of evaluation of scholarly communication data sources. I propose a methodology enabling semi-automatic, large scale disciplinary comparison of the scope and quality of bibliographic databases, which I used to evaluate a novel source of information.

I am continuing my efforts begun during master's work to quantitatively study the openness, quality, scope and transparency of one of the novel sources on scholarly communication, the Microsoft Academic [10]. The work has evolved into creation of a generalised, semi-automatic methodology enabling large-scale quantitative studies of a variety of scholarly communication data sources, such as bibliographic databases, institutional repositories, journals, or any of the innovative services arising. I have also used the designed methodology, based on disciplinary keywords sourced from review articles, for a comparison of Microsoft Academic, Google Scholar and Scopus - results of which will be published later this year.

4 A SOCIAL MACHINE FOR ACADEMICS' ONLINE PROFILES

In the next step, I plan to scrutinise the available online researchers profiles, by adapting the above-described methodology to quantitatively study the types of information presented and frequency of updating of researchers online profiles. This activity is going to uncover the current state and fragmentation of online academics' profiles. We have observed the growth of novel services such as ResearchGate, Academia.edu or ORCID - alongside with private and departmental websites. The detailed studies of the quality, timeliness, types and form of information presented by researchers about themselves and their work on such platforms, have only been conducted by a few researchers in a limited scope [7–9].

Finally, I aim to propose a social machine design - a system for online researchers' profiles using the scholarly communications data, computational methods, combined with human input. Outcomes of this activity and previous study will be verified in a questionnaire, asking scientists about the information on other researchers they are frequently looking for online, opportunities and barriers of existing online profiles, and their own needs for types of presented information.

REFERENCES

- [1] Bianca Amaro, Paola A Azrilevich, Argentina Dominique Babini, Argentina Gerald Beasley, Norbert Lossau, Katharina Mueller, Kostas Repanas, A * Star, Singapore Oya Rieger, Eloy Rodrigues, Coar / Openaire, Judy Ruttenberg, Marta Viragos, and Xiaolin Zhang. 2015. Promoting Open Knowledge and Open Science Report of the Current State of Repositories. (2015).
- [2] Tim Berners-Lee. 1989. Information Management: A Proposal. 2016, 17-09-2016 (1989). <https://www.w3.org/History/1989/proposal.html>
- [3] Lutz Bornmann. 2015. Usefulness of altmetrics for measuring the broader impact of research. *Aslib Journal of Information Management* 67, 3 (2015), 305–319. <https://doi.org/10.1108/AJIM-09-2014-0115> arXiv:1406.7611
- [4] Jody Condit Fagan. 2017. An Evidence-Based Review of Academic Web Search Engines, 2014-2016: Implications for librarians' practice and research agenda. *Information Technology and Libraries* 36, 2 (2017), 7–47. <https://doi.org/10.6017/ital.v36i2.9718>
- [5] Sven E. Hug and Martin P. Braendle. 2017. The coverage of Microsoft Academic: Analyzing the publication output of a university. *arXiv.org* (2017), 1–23. arXiv:1703.05539 <http://arxiv.org/abs/1703.05539>
- [6] JISC. 2012. *Researchers of Tomorrow: the research behaviour of Generation Y doctoral students*. Technical Report June.
- [7] X Li, Q Wu, and N Zhang. 2017. Citation personal display: A case study of personal websites by physicists in 11 well-known universities. *Journal of Documentation* 73, 4 (2017), 733–747. <https://doi.org/10.1108/JD-09-2016-0108>
- [8] Amalia Más-Bleda and Isidro F. Aguillo. 2013. Can a personal website be useful as an information source to assess individual scientists? The case of European highly cited researchers. *Scientometrics* 96, 1 (2013), 51–67. <https://doi.org/10.1007/s11192-013-0952-5>
- [9] José Luis Ortega. 2015. Disciplinary differences in the use of academic social networking sites. *Online Information Review* 39, 4 (2015), 520–536. <https://doi.org/10.1108/OIR-03-2015-0093> arXiv:arXiv:1011.1669v3
- [10] Bartosz Paszcza. 2016. *Comparison of the Microsoft Academic Graph with other scholarly citation databases*. Ph.D. Dissertation. University of Southampton. <https://doi.org/10.13140/RG.2.2.21858.94405>
- [11] Jason Priem, Dario Taraborelli, Paul Groth, and Cameron Neylon. 2010. Altmetrics: A manifesto. (2010). <http://altmetrics.org/manifesto/>
- [12] James Wilsdon. 2015. *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. Technical Report. HEFCE. <https://doi.org/10.13150/RG.2.1.4929.1363>

The digital divide in Amsterdam: the perspective and needs of families with a low socio-economic position

N.S.Goedhart, MSc.
Athena Institute for Research on
Innovation and Communication,
VU University
Amsterdam, the Netherlands
n.s.goedhart@vu.nl

R. Kattouw, MA
OOSTWEST
Amsterdam, the Netherlands
rolinka@oostwest.eu

Dr. C. Dedding
Athena Institute for Research on
Innovation and Communication,
VU University
Amsterdam, the Netherlands
c.dedding@vu.nl

Abstract— This explorative participatory study presents a probing portrait of the digital divide in Amsterdam. This study shows how complex digital inclusion is, in the context of poverty and low-literacy, and presents several steppingstones for action.

Keywords—digital divide, digital inclusion, poverty, participatory research

I. INTRODUCTION

Access to and mastery of ICT are increasingly indispensable participatory conditions in the Western society, e.g. to have access to education, to health care, to search for job possibilities, and to get in touch with governmental agencies. Also, the Netherlands facilitates more and more services online, while 17% of the Dutch citizens have no or low ICT skills (1). Socioeconomically disadvantaged citizens are more likely to be excluded in terms of motivational, material, skills and usage access to ICT (2–4). The relationship between digital and the broader persisting socioeconomic inequalities suggest that ICT can empower socioeconomically disadvantages citizens and has the potential to break the pattern of generative poverty (5).

In Amsterdam, a large part of the population is likely to fall behind in the digital society, as 18% of the citizens is low-literate and 24% lives in poverty (6). At present, little is known about the mechanisms explaining ICT (non-) use in everyday lives of families with a low socio-economic position (SEP); also little is known about the experiences, perspectives, and needs of people living in such circumstances. By means of a participatory qualitative study, we explored the barriers that families with a low SEP (children in the age of 8 to 15 and their parents) encounter in ICT use and their needs in order to develop policy and interventions that fit their needs and reality in our quickly growing digital society.

II. METHOD

From April 2017 to August 2017 data-collection took place. This participatory research gathered qualitative data through participant observations (n=13), semi-structured interviews (n=9), and dialog sessions (n=3). This approach was chosen because it helped to gain an understanding of views and personal experience of families with a low SEP and what they

believe needs to be done in order to increase inclusion. In total, we reached 51 mothers, 23 fathers, and 64 children (8-15 year). This study conceptually draws upon Van Dijk's multiple access model, acknowledging four successive types of access: motivation, material, skills, and usage access. Full transcripts and field notes were analyzed using ethnographic content analysis using Atlas.ti.

III. RESULTS

The parents are most often low educated, low-literate and usually experience problems in the field of housing, work, and income. The daily struggles, e.g. managing sporadic income, making difficult economic trade-offs or dealing with low linguistic abilities, has an enormous influence on ICT (non-) use. For example, Jantien¹ explains that she needs to go to the public library to print homework for her daughter. Also for the wrong pages, she needs to pay, sometimes this mistake cost her up to €2,20. She can also buy a loaf of bread for that money.

The fathers and mothers are aware of the importance of having access to the online world for their children, but also for themselves. Without a computer or the skills to use it, they feel handicapped. The rapidly digitalizing world lead to a feeling of exclusion.

"I saw a job offer. [...] something I could do. So I went to that office with my CV and motivation letter. Arriving there the man behind the desk only explains to me that I need to send my CV by email. I can't. And then I come home and think: screw it, alright." (Carl)

"Sometimes I think where can I get that information? One day I went to the town hall. The only thing I got there was a piece of paper with instructions to log in to another complicated website." (Orlando)

Not being involved in the online world does not only impact participation in the society, but also the relationship between parents and their children. For example, Kemal (12 years) explains that his mother is often *"Angry with the computer."* He needs to help his mother with filling in forms or checking her mail or bank account. Meaning that children get involved in adult businesses which aren't meant for them.

¹ To guarantee anonymity, fictitious names are used in the results section

Barriers

Barriers to using ICT are related to all four successive types of access as described by van Dijk. The results related to motivation indicate that the group who *needs* to use ICT is much larger than the group who *wants* to learn ICT skills. The group who *needs* to use ICT often visits social services to work something out, but are, by circumstances, 'not motivated' to learn ICT. They prioritize taking care of their children, managing their household and making ends meet every week, often bothered by feelings of guilt and shame for not being able to provide for their children.

Related to material access our findings show that most households have access to at least one ICT device: a mobile phone, laptop, tablet or computer. However, this does not automatically mean that these devices work properly, nor that all families have access to the online world. Basic conditions, e.g. Wi-Fi, up to date software, or printers, are often missing. Fayiza highlights: "*Wi-Fi is more important than a computer. A computer without Wi-Fi is useless.*". Furthermore, most families depend on cheap or second handed material that is often broken and not user-friendly.

Participant observations show that the skills of both parents, as well as their children, is limited to so-called splinter skills. Mothers who seem quite skilled become stressed when they accidentally push the Caps-lock bottom and children who can perfectly play Minecraft have no idea how to save a Word document. Splinter skills cause difficulties and frustrations. Moreover, they gloss over what people are actually capable of.

"I have an email address. [...] I use it if my doctor asks me to give my email. If someone sends me an email I can react on that, but I can't make a new one." (Doria)

However, many parents indicate that illiteracy is the main issue that restricts them in using ICT; language is essential in understanding the often complex websites.

YouTube is an interesting exception, it is the most often used program to find information about cooking, health, household chores, etc. Also, Google, WhatsApp, and Skype are often mentioned. Besides their user-friendliness, mothers can use their first language to use these programs.

Needs

Asked for their needs, parents often refer the needs of their children. Nilofar, explains: "*I want a computer for my children. They get to know all kind of websites on school and TV. I want my children to be happy.*". Mothers also explain that there is a high need for homework assistance. Children living in families with a low SEP depend on informal, free after-school childcare. Observations show that in these informal care settings sometimes even simple playing and learning materials, e.g. crayons, paper, scissors, are missing. Let alone, that children are stimulated to learn more about ICT and the online world.

For most mothers, the digital world lacks a friendly, trusted face. Lyobosa describes: "*Just having a computer doesn't make sense*". Mothers like to have a place close to their homes where they can learn ICT. This place should be without daily - family - distractions; where they feel comfortable and truly welcome. Often having experiences of failures and fines, they prefer

personal attention from a teacher they can trust. A 'walk-in-hour format', matches their practical and personal needs; they want to learn things they need to know at that specific moment.

IV. DISCUSSION: HOW TO PROCEED?

This study shows how urgent, but at the same time how complex digital inclusion is in the context of poverty. We should search for creative out-of-the-box interventions to increase proper working hard- and software and to educate families without or with low ICT skills. Interventions should match parents' practical and personal needs with respect to the complex daily reality in which people live (7-8). Interventions need to be efficient in the sense of time, motivation and energy since people have other important projects to deal with. Also, there is a need to intervene early, to develop easily accessible places where children can be creative, explore and experiment with ICT in a playful way (4).

Interestingly, YouTube, Google, and Whats-App manage to reach people with low resources and low ICT skills. Unfortunately, the basic principles of those businesses are still rarely or not used by official authorities in the Netherlands. The websites often contain an overload and too difficult language and are not as user-friendly. Our results suggest that it seems more promising to change websites to the capabilities and strengths of the most vulnerable groups than to teach this group how to use ICT in a world that has not been designed for them. We are aware that this brings along legal challenges, but we could at least explore the opportunities.

This participatory qualitative study will be followed up by interventions developed with and for families with a low-SEP in Amsterdam. A broad coalition is needed in which businesses, creative industry, policy makers, knowledge institutions, volunteers, and citizens work together. Interventions should be developed with a view to promoting ICT skills of the most vulnerable group, but also to design the online world to the capabilities and strengths of the most vulnerable group. The question is who takes up the responsibility, but also what and how to prioritize.

REFERENCES

- [1] CBS. (2017). *ICT, kennis en economie 2017*. Den Haag.
- [2] Fuchs, C. (2009). The role of income inequality in a multivariate cross-national analysis of the digital divide. *Social Science Computer Review*, 27(1), 41-58.
- [3] Helsper, E. J., & Reisdorf, B. C. (2017). The emergence of a "digital underclass" in Great Britain and Sweden: Changing reasons for digital exclusion. *New media & society*, 19(8), 1253-1270.
- [4] Helsper, E.J., van Deursen, A.J.A.M. & Eynon, R. (2015). *Tangible Outcomes of Internet Use*. From Digital Skills to Tangible Outcomes project report. Available at: www.oii.ox.ac.uk/research/projects/?id=112
- [5] Dedding, C., Goedhart, N.S., & Kattouw R. (2018). Digitale ongelijkheid, een participatieve verkenning in Amsterdam. Available at: http://www.pact-amsterdam.nl/wp-content/uploads/2018/01/Bouwstenen-voor-Digitale-Inclusie_Compleet.pdf
- [6] Michon, L., Holaind, N., Slot, J., Nottelman, N., & Wenneker, C. (2017) *Amsterdamse Armoedemonitor 2016*. Amsterdam.
- [7] Grunwald, K., & Thiersch, H. (2009). The concept of the 'lifeworld orientation' for social work and social care. *Journal of Social Work Practice*, 23(2), 131-146.
- [8] Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. *science*, 341(6149), 976-980.

Work in progress, don't cite

Cyber Hate against immigrants: an Italian case

Komal Florio

Università Degli Studi di Torino, Italy

Viviana Patti

Università Degli Studi di Torino, Italy

ACM Reference Format:

Komal Florio and Viviana Patti. 2018. Cyber Hate against immigrants: an Italian case. In *Proceedings of WebScience 2018 (WebSci18)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 INTRODUCTION

The presence and integration of immigrants is one of the most controversial issues in our society, and given current worldwide political instabilities, it will likely become ever more prominent in cultural and political debates. Social media play an increasingly important role in how citizens debate opinions and react to events with (perceived) local impact, with online platforms frequently providing a channel for hate speech ('cyberhate'). We propose here a novel approach based on the integration of fine-grain sentiment analysis on cyberhate against this target in the Italian Twittersphere with offline indicators, provided by the main suppliers of official statistical data (e.g., the Italian National Institute of Statistics, IS-TAT)), such as indicators of quality of life (e.g. unemployment rate, education level, proportion of resident immigrants, presence of CIE¹) with the final aim to better understand the cultural and social backgrounds in which this sort of aggressive responses originates².

2 METHODOLOGY

Our ultimate and innovative goal is to infer a detailed description of the social context in which cyberhate against immigrants originates in Italy by performing a fine grained sentiment analysis on the contents of the Italian Twittersphere which are related to this topic and combine these results with more traditional statistical indicators about immigration in general. In order to correctly interpret the results we will find, we think the first step of our approach consists in a detailed discussion about the representativeness of our datasets and how this can affect any claims we will be able to make. In this context, a fundamental aspect of our research will be to assess the representativeness of Twitter users demographics in Italy and the relation with the digital divide in Italy.

Characterization of the Italian Twittersphere. Twitter seems the ideal social media for opinion mining in polarized debates, due to accessibility of its users activity and richness of data about contents, contents producer and their interactions. Nevertheless only a fraction of the Italian population regularly active on the Internet engage in this sort of activity, hence we think an assessment of the

¹Centers for identification of refugees

²For a wider context of the work see <http://hatespeech.di.unito.it/>.

Permission to make digital or hard copies of part or all of this work for personal or internal use, or the internal or personal use of specific clients, is granted by ACM for users registered with ACM, provided that the fee of \$12.00 is paid directly to ACM. This fee code for users registered with ACM: 978-1-4503-6666-6/18/0000-0000. Distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
WebSci18, May 2018, Amsterdam, Netherlands
© 2018 Copyright held by the owner/author(s).
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.
<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

2018-04-05 21:09 page 1 (pp. 1-2)

number of active users and a detailed analysis of direct and inferred demographic data (as in [1, 6]) are essential in order to estimate the size and representativeness of this phenomena. Moreover, recent literature in sociology shows that *communities* play an important role in polarized debates, hence we propose to perform a network analysis of the users in our dataset in order to better capture and describe this dynamic and add an extra dimension to the cyberhate phenomena [3, 4, 7].

Digital Divide. The access to online services is, in general, not uniformly distributed across a country: due to the so called 'digital divide' and Internet penetration rate, data related to people living in areas with good infrastructure may be over-represented. We will characterize the "digital divide" in Italy using data from the Italian Communication Authority³ and we will try and detect interesting patterns (e.g.: urban VS rural areas). In general, we cannot thus consider results from online sources to be representative for the whole population of a Country, but only for specific geographic locations: nevertheless it will be interesting to explore any local geographical patterns in the diffusion of hateful opinions across regions and correlate this results with statistical indicators about local characterization of the immigration phenomena.

Data collection strategies. For what concerns data collection, our research relies on the TWITA⁴ project dataset: a collection of tweets identified as being written in Italian. This dataset was then filtered for our purpose by selecting tweets with a geotag in Italy and containing one of the following keywords⁵: *immigrat** (immigrants), *immigrazione* (immigration), *migrant** (migrant), *profug** (refugee), *stranier** (foreigner), *terrorismo* (terrorism), *terrorist** (terrorist), *islam*, *corano* (Koran), *musulman** (Muslim), *rom* (Roma), *nomad** (nomad). This set of keywords include neutral terms related to immigrants in order to collect as many tweets as possible on the topic, but purposely exclude negative epithets in order to capture different possible shades of sentiment in the online discourse on this topic (e.g.: offensiveness, hostility,...). We decided to include the term "terrorism/terrorist" which has a negative connotation but it cannot strictly be categorized as an insult. Due to historic reasons of how immigration has happened in Italy⁶, cyberhate has assumed different connotation depending on the targets, hence we decided to include specific terms to capture this fine grained distinction regarding a specific religion (which is a more recent phenomena) and Roma ethnicity (victims of racism or hateful behaviors that mainly stem from prejudices).

Bots detection. Bots have been playing an increasing and much discussed roles on online debates [2, 8]: we aim to characterize their influence in the context of Italian Twittersphere.

³AGCOM

⁴Twita project: <https://valeriobasile.github.io/twita/about.html>

⁵The * stands for the different ending of words in Italian indicating the masculine, feminine and plural.

⁶"Jo Cox" Commission report on intolerance, xenophobia, racism and hate phenomena on Italy: <http://www.camera.it/leg17/1313>

3 CROWDSOURCED ANNOTATION EXPERIMENT: PRELIMINARY RESULTS

We run a first crowdsourced annotation experiment using Figure Eight platform ⁷ with 2,200 tweets from the dataset filtered as described before. The detailed description of the annotation scheme applied is provided in [5]. In order to capture the different shades of cyberhate, each tweets was annotated not only with the presence of hate-speech (HS) but also for its *intensity* and the presence of other features: *offensiveness*, *aggressiveness*, *irony* and *stereotype*. To start with we choose tweets from the first six months of 2015: the most recent year for which we have complete datasets from the ISTAT ⁸ about immigrants in Italy and thus allowing us to integrate online data with more traditional offline statistical indicators. We aim to run comparative results on different time frames in order to give a temporal perspective of our finding.

HS		intensity			
no	yes	1	2	3	4.0
84.5	15.5	36.4	36.7	4.2	3.3

		intensity			
		1	2	3	4
aggressiveness	no	62.5	53.9	31.3	13.8
	weak	16.0	40.8	31.2	17.8
stereotype	no	39.6	43.9	37.5	41.2
	yes	74.2	56.1	32.5	58.8
offensiveness	no	18.0	35.8	25.0	0.0
	weak	77.3	53.5	18.8	29.4
irony	no	80.4	44.8	8.2	7.7
	yes	19.6	13.9	0.0	1.0

Figure 1: Hate speech and intensity labels distribution in percentage

15.5% of tweets were classified as HS with different intensities. Analyzing our label distribution we found that offensiveness co-occurs with HS in 78% of the cases, while stereotype in 70% of the cases. The distribution for the other labels are coherent with the findings in [5]. Interestingly, we can see from the table on the bottom and from Fig. 2 that aggressiveness and offensiveness correlate positively with intensity, while irony correlates negatively. In general, high values of intensity co-occur with strong aggressiveness and offensiveness, therefore such dimensions appear to be promising indicators in detecting tweets with high intensity of HS, which are potentially the most dangerous ones, that is the ones that could be urgent to detect in order to prevent violent actions. This preliminary results are in substantial agreement with the finding of the similar experiment run on data from 2016-2017, and described in details in [5].

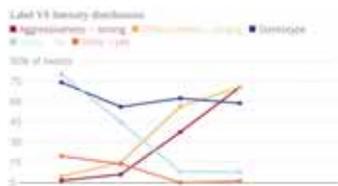


Figure 2: Trends in intensity VS other labels

⁷<https://www.figure-eight.com/> (formerly Crowdfower)

⁸<http://www.istat.it>

Tweet	Traduzione
«Grazie mille per eliminare tutti noi e pulirci dalle loro e mandarci il buio è luce»	«Grazie mille per eliminare voi e noi e darci luce. Buio e luce e noi e loro»
«Perché la soluzione immigrazione è fare come fanno in Spagna, un voto lì punto per candidato»	«Perché? la soluzione è immigrazione è fare come fanno in Spain, 100 gr of lead for each illegal immigrant»
«Mandare tutti i loro bambini uccidendo gli bambini possibile, perché i bambini e il legionario la gente per lavoro è evolvibile»	«Bomb of atomic countries killing as many children as possible, because they will grow up and sell us. Parents to cover on it is evolvibile»

Figure 3: Samples from most dangerous tweets

We show in Fig. 3 a sample of the most dangerous tweets according to our annotation (strong aggressiveness and offensiveness, intensity 4, stereotype yes). Interestingly, some display hate incitement towards a group or religion combined with misogyny towards an individual: in particular Italian politician Laura Boldrini, former President of the Chamber of Deputies of Italy, who has been targeted with a significant amount of HS messages in last years.

4 CONCLUSIONS AND FUTURE WORKS

We believe that our work will lead to interesting finding due to the innovative approach of combining sentiment analysis with traditional statistical indexes with the aim of describing the social context in which cyberhate against immigrants in Italy originates. We show that the multidimensional annotation we applied can highlight interesting features to predict the most dangerous tweets, hence we think that such deeper corpus analysis will be used to implement an automatic classifier that will enable us to analyze larger datasets. Future perspectives of this work include also a geographical mapping of our results combined with more traditional statistical indicators about immigration in Italy. I think my work would benefit greatly from attending WebSci2018 because some of the events are dedicated to crucial issues in our research (such as digital divide, interdisciplinary research, digital sociology, online information quality, source criticism beyond bots). Moreover the interaction with speakers and attendees will be meaningful in terms of ideas exchange, possible collaborations and mentoring.

REFERENCES

- [1] Pablo Barberá and Gonzalo Rivero. 2015. Understanding the Political Representativeness of Twitter Users. *Soc. Sci. Comput. Rev.* 33, 6 (dec 2015), 712–729. <https://doi.org/10.1177/0894439314558836>
- [2] Zafar Gilani, Reza Farahbakhsh, and Jon Crowcroft. 2017. Do Bots impact Twitter activity?. In *Proceedings of the 26th International Conference on World Wide Web Companion*. International World Wide Web Conferences Steering Committee, 781–782.
- [3] Mirko Lai, Marcella Tambuscio, Viviana Patti, Giancarlo Ruffo, and Paolo Rosso. 2017. Extracting Graph Topological Information and Users' Opinion. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction*. Springer International Publishing, 112–118.
- [4] Yiye Ruan, David Fuhry, and Srinivasan Parthasarathy. 2013. Efficient community detection in large networks using content and links. In *Proc. 22nd Int. Conf. World Wide Web - WWW '13*. ACM Press, New York, USA, 1089–1098. <https://doi.org/10.1145/2488388.2488483>
- [5] Manuela Sanguinetti, Fabio Poletto, Cristina Bosco, Viviana Patti, and Marco Stranisci. 2018. An Italian Twitter corpus of hate Speech against Immigrants. *LREC proceedings, to appear* (2018).
- [6] Luke Sloan, Jeffrey Morgan, Pete Burnap, and Matthew Williams. 2015. Who Tweets? Deriving the Demographic Characteristics of Age, Occupation and Social Class from Twitter User Meta-Data. *PLoS One* 10, 3 (mar 2015), e0115545. <https://doi.org/10.1371/journal.pone.0115545>
- [7] Stefan Stieglitz and Linh Dang-Xuan. 2013. Social media and political communication: a social media analytics framework. *Soc. Netw. Anal. Min.* 3, 4 (dec 2013), 1277–1291. <https://doi.org/10.1007/s13278-012-0079-3>
- [8] Onur Varol, Emilio Ferrara, Clayton A Davis, Filippo Menczer, and Alessandro Flammini. 2017. Online human-bot interactions: Detection, estimation, and characterization. *arXiv preprint arXiv:1703.03107* (2017).

An explanation of the changing relations between digital platforms and entrepreneurship in Trinidad and Tobago

Keisha Candice Taylor, Dr. David Baxter, Prof. Peter Sunley

Southampton Business School, University of Southampton, Southampton, SO17 1BJ, UK,
keisha.taylor@soton.ac.uk d.baxter@soton.ac.uk p.j.sunley@soton.ac.uk

Abstract

Digital platforms are increasingly influencing individual entrepreneurship, their processes and their ecosystems but their relationship is insufficiently discussed in the entrepreneurship literature. The thesis addresses this gap by examining the relationship between digital platforms and individual entrepreneurship in Trinidad and Tobago, where despite high levels of internet and platform usage, its influence remains mostly unacknowledged in entrepreneurship research.

CCS concepts

Human Computer Interaction

Keywords

Entrepreneurship, Digital Platforms
Technological Affordances and
Constraints

1) Introduction

Social and economic life are becoming more digitised [1-3]. This is in part driven by the use of digital platforms. However, the relationship between individual entrepreneurship and digital platforms has been somewhat lost in the management and entrepreneurship literature [1, 3]. To understand this relationship the thesis examines the case of Trinidad and Tobago, a twin island, multi-cultural oil and gas economy that is currently in recession. While the country has high levels of internet access and platforms are widely used the influence of digital platforms has been largely left out in research about entrepreneurship in the country. The primary source of information on individual entrepreneurship for Trinidad and Tobago comes from the Global Entrepreneurship Monitor (GEM), which

supports local research. GEM's research in 105 countries thus far is highly cited by governments, media and academia [4], but does not consider the impact of digitisation on entrepreneurship. The types of platforms under study include social media platforms (for example Facebook). Research revealed that the much used Facebook wields considerable influence in the life of Trinidad and Tobago citizens [5]. E-commerce platforms like Amazon will also be reviewed. An estimated 1 billion TT dollars worth of goods are bought online through platforms like Amazon and other online stores. This has contributed to a foreign exchange crisis and the subsequent introduction of 7% tax on all goods purchased overseas online [6]. Newer on-demand platforms (for example, Uber) will also be studied. In 2017 Uber began operations in Trinidad and Tobago, a country with a tradition of unlicensed taxis use [7]. The objective of the research is therefore to utilise TACT to understand the relationship between digital platforms and entrepreneurship in Trinidad and Tobago. Through in-depth examination, it also aims to provide policy recommendations to the Trinidad and Tobago government and other countries with similar characteristics.

2) Individual Entrepreneurship

The definition of entrepreneurship that is most suitable is that of GEM, which has been narrowed to focus on the individual. Entrepreneurship is therefore defined as 'any attempt at new business or new venture creation, such as self-employment, a new business organization, or the expansion of an existing business, by an individual' [8, p. 3]. The entrepreneurship process is usually viewed as lineal though the use of digital platforms may render the

process non-linear [3]. Therefore 5 phases of entrepreneurship as employed by GEM will be integrated in this analysis. These will be explored in the context of the social, economic, political and cultural environment within which it takes place.

3) Digital platforms

Digital platforms are defined as digital frameworks that support collaboration, the provision and exchange of information, products and/or services among users by connecting two or more individuals or participant groups or varying sides of a market [9]. While not all digital platforms utilised by entrepreneurs will be discussed, the research will help to identify platforms that are most relevant to entrepreneurs in Trinidad and Tobago. Investigation of multiple types of platforms is valuable as entrepreneurs may not use platforms in ways designers intended.

4) Technology Affordances and Constraints Theory

To examine this relationship, the thesis uses Technology and Affordances and Constraints theory (TACT) According to TACT [3, 10-12], technological affordances refer to the various ways that a particular technology, can potentially be used. Constraints, which are equally important are obstacles that technology may present at the same time [13]. This theory recognises that technology is not always helpful and may have negative or unintended consequences. This is particularly important given contentious debates on the extent to which digital platforms democratise or dominate [9, 14].

5) Research Design

A multi-method qualitative research design is adopted. This includes interviews, 3 focus groups and the review of secondary data, which includes relevant government reports and platforms used by entrepreneurs. It is hoped that this research would help to improve explanations of the relationship between digital platforms and entrepreneurship by providing new insight from Trinidad and Tobago.

<REF>References

- [1] S. Nambisan, K. Lyytinen, A. Majchrzak, and M. Song, "Digital Innovation Management: Reinventing Innovation Management Research in a Digital World " *MIS Quarterly*, vol. 41, pp. 223-238, 2017.
- [2] Z. J. Acs and F. Sussan, "The Digital Entrepreneurial Ecosystem " *Small Business Economics*, vol. 49, pp. 55-73, 2017.
- [3] E. Autio, S. Nambisan, L. Thomas, W. D, and M. Wright, "Digital Affordances, Spatial Affordances, and the Genesis Of Entrepreneurial Ecosystems," *Forthcoming in the Strategic Entrepreneurship Journal*, 2017.
- [4] N. Bosma, *The Global Entrepreneurship Monitor (GEM) and Its Impact on Entrepreneurship Research*: Boston and Delft: Now; distributed by Lightning Source, La Vergne, Tenn., 2013.
- [5] D. Miller, E. Costa, N. Haynes, T. McDonald, R. Nicolescu, J. Sinanan, *et al.*, *How the world changed social media*: UCL press, 2016.
- [6] Oxford Business Group, "The Report: Trinidad and Tobago," Oxford Business Group 2017.
- [7] R. John-Lall. (2017, 1 August 2017). *Uber is here!* Available: <http://www.guardian.co.tt/business-guardian/2017-01-19/uber-here>
- [8] P. D. Reynolds, M. Hay, and S. M. Camp, " Global Entrepreneurship Monitor: 1999 Executive Report," Kauffman Foundation., Kansas City, MO 1999.
- [9] A. Hagiu and J. Wright, "Multi-sided platforms," *International Journal of Industrial Organization*, vol. 43, pp. 162-174, 2015.
- [10] J. J. Gibson, Ed., *The theory of affordances* (Perceiving, Acting and Knowing. Hillsdale (NJ): Lawrence Erlbaum Associates., 1977, p.^pp. Pages.
- [11] P. M. Leonardi, "When flexible routines meet flexible technologies: Affordance, constraint, and the imbrication of human and material agencies," *Management Information Systems Quarterly*, vol. 35, pp. 147-167, 2011.
- [12] A. Majchrzak, L. M. Markus, and J. Wareham, "Designing for Digital Transformation: Lessons for Information Systems Research from the Study of ICT and Societal Challenges " *Mis Quarterly*, vol. 4, pp. 267-278, 2016.
- [13] O. Volkoff and D. M. Strong, "Critical realism and affordances: theorizing IT-associated organizational change processes," *Management Information Systems Quarterly*, vol. 37, pp. 819-834, 2013.
- [14] J. A. Schwarz, "Platform Logic: An interdisciplinary approach to the platform-based economy," *Policy and Internet*, vol. 9999, 2017.

APTENDOFDOCAPT

Crowdsourcing Application for Responsible Production in Africa

Promoting responsible and sustainable business through crowd-generated transparency

André Baart

Amsterdam Business School, University of Amsterdam

Amsterdam, the Netherlands

a.h.j.baart@uva.nl

ABSTRACT

We propose a case study on the development of a crowd sourcing platform that promotes responsible production in the context of the developing world. By leveraging the “Wisdom of the Crowd”: by developing a platform that allows the contribution of data by all involved parties and local experts in the production of products and resources, making this data available to the public and promoting constructive dialogue between parties; Best practices and patterns of responsible production can be researched and promoted in order to improve livelihoods.

KEYWORDS

ICT4D, crowdsourcing, wisdom of the crowd, sustainable production, responsible production, conflict minerals, mobile applications, transparency, human rights

1 INTRODUCTION

The products and goods that are available on the (international) market have become increasingly more complex, and often depend on resources that are not available in the country of the consumer. The production chains that allow for the production of these goods are often long and invisible to the consumer. The resources in these chains come from all over the world and often originate from developing countries, where the production conditions are often harmful for the workers and the environment. While there are many initiatives in terms of the creation of certification schemes and organizations, there is still a lack of transparency and monitoring of these conditions, often caused by corruption and a lack of press freedom.

An example of a resource that has been getting a lot of attention recently is cobalt. Cobalt is a mineral that is used in the production of lithium-ion batteries. The rise in sales of electric vehicles has caused a significant rise in the price of cobalt, and the global demand is expected to nearly triple in the next 12 years.¹ More than half of the world’s cobalt is produced in the Democratic Republic of the Congo (DRC), where the mining industry is characterized by human rights violations, corruption, armed conflict,

unsafe working conditions and severe environmental damage. Furthermore the DRC and other sub-Saharan African countries have weak governments, that have a low enforcement capacity. The lack of information from these regions prevent firms and consumers from verifying the local production conditions. These problems have caused firms –pressured by consumers demanding sustainable production practices– to stop buying resources from these countries as long as conditions are bad or unknown, or to set up new certification schemes to ensure customers that products are “conflict-free”.

The economies of many developing countries (including DRC) are dependent on the export of resources the like of cobalt, therefore a decline in exports might cause further loss of livelihoods. [5] Furthermore, the low availability of responsible resources cause an increase of the prices of many products on the international market. This could delay the adoption of technologies that are essential in the transition to a more sustainable way of living –such as the adoption of electric vehicles– which are essential to achieving the UN’s Sustainable Development Goals and the Paris climate agreement.

The recent rise in availability and uptake of mobile telephony and internet in developing countries, provides new opportunities in gathering local information on the conditions in production chains. By developing a platform that facilitates the reporting and monitoring of incidents, as well as initiatives undertaken to improve production conditions, the locally (but dispersed) available knowledge and information can be aggregated and made available to the public. The platform should involve all relevant parties in production chains such as: production companies, local NGOs, human rights lawyers, activists, MNEs, unions, policy makers and government; A constructive dialogue between these parties can be promoted through the platform, while gathering data that allows research on the best practices in responsible production by systematic analysis of this data. By aggregating the inputs of all involved parties a representative overview of the local situation can be formed, the “Wisdom of the Crowd”. [4] The data (anonymized if necessary) will be made available to the public, allowing academia and consumers to research the issues in the production of resources and commodities, aiding policy makers and MNEs in ensuring the improvement of the sustainability of production chains.

While there are several theories on the factors that play a role in setting up successful crowd sourcing initiatives, there is little research on applying these theories in the context of the developing world. [2, 3] A framework is needed that can be used to facilitate the formation of crowd sourcing initiatives in the developing world, ensuring significant uptake, data reliability, incentives and infrastructure.

¹ See: <https://www.bloomberg.com/news/articles/2018-03-27/cobalt-upstarts-add-processing-capacity-in-congo-as-prices-surge>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

WebSci’18, May 2018, Amsterdam, the Netherlands

© 2018 Copyright held by the owner/author(s).

ACM ISBN 123-4567-24-567/08/06.

https://doi.org/10.475/123_4

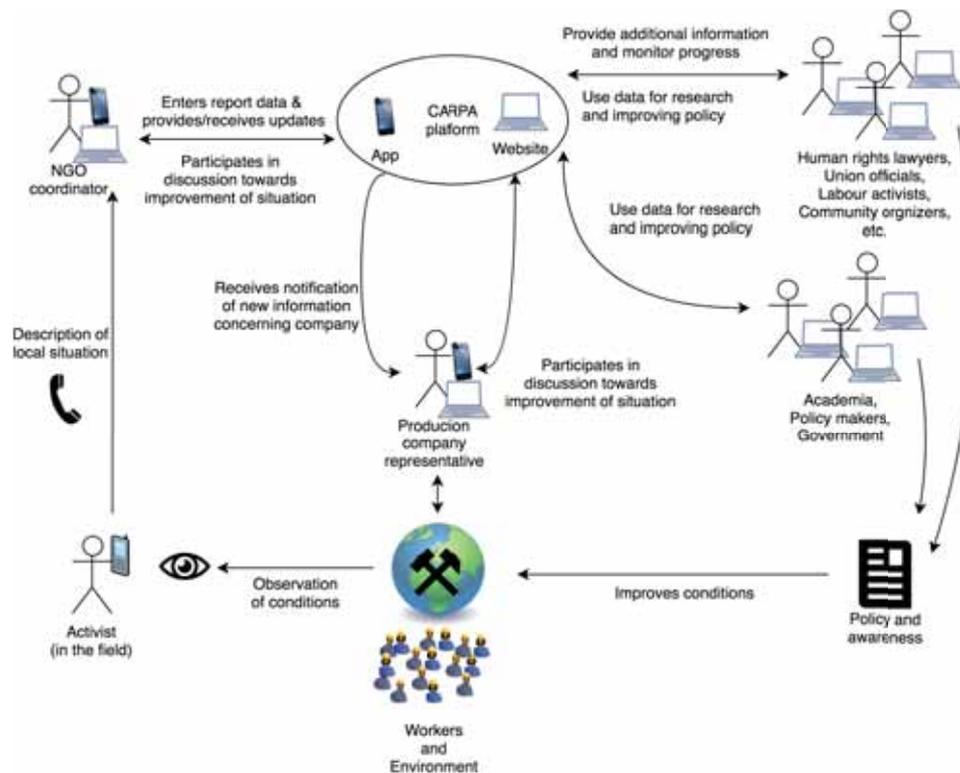


Figure 1: General overview of the CARPA crowd sourcing platform.

2 METHODOLOGY AND APPROACH

The approach to developing a framework for the application of crowd sourcing and Wisdom of the Crowd theories in the developing context, will be a case study in the DRC, Mali and South-Africa. (countries with significant mining production and exports) We are in the process of identifying an initial set of *lead users* for the platform, in cooperation with local partners that participate in the research. These users will be selected based on their potential for the platform, which is determined by (but not limited to) their local network, integrity, reliability, technical ability and openness to using 'new' technologies. After selecting and contacting these parties, the requirements of the platform will be determined by analyzing the existing flows of information within these organizations and structures and determining what data is to be gathered for achieving the goals of the platform. (improving sustainability and responsibility through policy, academic research, etc.) After the initial requirements have been determined, the platform (consisting of a mobile application and a website) will be developed using the *agile* development methodology. Iterative development is especially suited in the developing context, as to-be users often have limited experience with using (complex) information systems, in which iterative prototype development can help in determining the desired functionality of the system. Furthermore the context of the developing world often results in requirements that are not often found in development projects oriented at the Global North, e.g. complete offline functionality, support for very old hardware.

[1] After the platform has matured in cooperation with the lead users, the user base of the platform will have to grow in order to improve the representativeness of the gathered data. By inviting more types of parties to join the platform, attracting media attention and iteratively expanding and improving the functionalities of the platform, we will gather data on responsible production as well as gain insight in how crowd sourcing can be applied to improve living conditions in the developing world.

In Figure 1 a general overview of the to-be-developed crowd sourcing platform is shown, indicating the involved parties and their roles in the platform.

REFERENCES

- [1] Anna Bon, Hans Akkermans, and Jaap Gordijn. 2016. Developing ICT Services in a Low-Resource Development Context. *Complex Systems Informatics and Modeling Quarterly* 9 (2016), 84–109.
- [2] Karthika Muthukumaraswamy. 2010. When the media meet crowds of wisdom: How journalists are tapping into audience expertise and manpower for the processes of newsgathering. *Journalism practice* 4, 1 (2010), 48–65.
- [3] Jay Pedersen, David Kocsis, Abhishek Tripathi, Alvin Tarrell, Aruna Weerakoon, Nargess Tahmasbi, Jie Xiong, Wei Deng, Onook Oh, and Gert-Jan de Vreede. 2013. Conceptual foundations of crowdsourcing: A review of IS research. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on*. IEEE, 579–588.
- [4] James Surowiecki. 2005. *The wisdom of crowds*. Anchor.
- [5] Ian Taylor. 2016. Dependency redux: Why Africa is not rising. *Review of African Political Economy* 43, 147 (2016), 8–25.

Online participation processes in the Green Party Germany

A qualitative and quantitative analysis of perceptions, expectations and experiences

Gefion Thuermer

Social Sciences, University of Southampton, UK
gefion.thuermer@soton.ac.uk

Supervised by: Silke Roth, Kieron O'Hara, and Steffen Staab

Project Summary

This project investigates the effect of the introduction of online participation processes in the Green Party Germany. Using a mixed methods approach, members' participation behaviour and perception thereof is evaluated over a period of two years, during which the processes are changed, with a view to broaden participation.

Aims

The project aims to understand the effect of the introduction of online participation on the participation behaviour of party members. Through a mixed methods approach it investigates how members change their behaviour when new online participation methods are introduced, or existing methods are changed. The goal is to find out

- How party members react to the introduction of new online participation processes and tools,
- Whether ideology influences the adoption of online participation, and
- Whether there are differences in views and tool adoption between party base and elite.

The study is based in political sociology. It uses Actor-Network Theory to map out the case and actors engaged in the processes, and mobilisation and reinforcement theory to evaluate the effects on participation.

Background

The Green Party Germany was founded in 1980, out of the women's, peace and environmental movements. Members and party leadership hold strong grass-roots democratic values, and participation is at the core of the party identity [1], [2]. This makes them an ideal case to investigate the effects of participatory processes. The party leadership specifically wants to engage as many members as possible in their decisions [3]. They believe that the internet can help them achieve this goal, which is why they introduce several online processes:

- the 'Mitgliederbefragung', introduced in 2016: an (at least) annual online survey among members, intended to bring their opinions into discussion processes in the parties' bodies and task forces;
- the 'Mitgliederbegehren', introduced in 2018: a petition system, by which a group of members can collectively 'demand' something from the executive board, or other groups in the party.
- the process by which members can submit proposals to delegate assemblies in 2016, to which a mandatory online verification for supporters was added. Twenty supporters are

required for all individual proposals, which used to be done either manually (signatures on paper) or en-bulk (one person collated a list of all supporters). With the new process, all members wishing to submit or support proposals need to either sign in to an online platform, or provide a signature.

However, although the goal is to reach more members, they are also very keen to not violate expectations of equal participation. Due to their roots in the women's movement, this refers particularly to gender equality – a strongly enforced focus across all levels and existing offline processes.

The digital divide literature suggests that the introduction of online processes, though indeed likely to reach more members, particularly in younger strata [4], is unlikely to increase participation *equally*. Differences in internet use with regards to gender, education and wealth persist [5]. Internet use for political participation is affected even more, and typically dominated by educated young men [5]. Moreover, the benefits derived from web use are influenced by other social inequalities [6]. The internet may reach a different target group, but it is just as much subject to participation biases as offline processes.

Two theories are commonly used to describe the effect of online participation. Mobilization theory suggests that new online opportunities to participate will engage a broader group [7], while reinforcement theory assumes that these will primarily be used by those who are already engaged, thus reinforcing existing advantages [8]. Vissers & Stolle [9] argue that these effects are not mutually exclusive, but can occur simultaneously for different groups. I propose to consider two additional potential effects: Replacement – participants move to online processes for things that they have previously done offline [4] – and non-use [10] – when online processes are simply not taken up at all. All of these have been investigated separately or in pairs – for example, by Kerr & Waddington [11] in a union context, or by Gerl et al. [12] for a smaller scale online process in the Green Party. However, these four effects have so far not been considered in combination, or in a single model.

Method

The methods used to conduct this study are both qualitative and quantitative, and applied in two phases. First, a mix of observations and interviews was used to evaluate the status quo of participation at the beginning of the project. This was focused on the decisions about the above described new online processes, and the discussion around those decisions. Participants were asked about their views on participation, online participation, and the new processes, to create a baseline for the analysis.

Based on the results, surveys were designed to assess the actual, planned, and likely future use of the new processes. Four

surveys are conducted in total: One among the entire member base, one only among users of the online verification system, and two panel surveys with a stratified sample of 4,236 members who are or are not communicating online with the party. The panel surveys in particular aim to measure the changes in behaviour over a period of time, during which the petition system is introduced, and use of the verification system and surveys is increased. All survey data is analysed using logistic regression models to evaluate which factors predict process use and changes in behaviour.

In addition to the surveys, observations, interviews and focus groups are conducted with members at different levels and engaged with these processes in different ways. This data is used to understand the perceptions of potential problems and benefits of these processes, as well as the considerations made during their implementation. All qualitative data is coded both inductively, to identify themes emerging from the data, and deductively, for evidence of the effects described in literature.

Findings

The first phase of the project has revealed that hopes and expectations of both party members and leadership are indeed very high: Most participants assumed that online processes would benefit other party members, but not be of particular benefit to themselves [13]. This is even more interesting in comparison to the first survey, which revealed that those who intend to use the new processes are mainly those who are already active, and beneficiaries of the digital divide [14]. Results suggest that there is ample evidence for reinforcement, in that members who already have an advantage are more likely to increase their participation. In fact, the new processes might undermine efforts of the party to achieve gender balance, if online processes do not replicate the strong measures they use offline. However, the most important piece of evidence – the second panel wave, which will allow measurement of changes to actual rather than anticipated participation – is still outstanding.

Benefits

This research will benefit the Green Party Germany specifically, and other parties by extension. The Green Party can be considered as a ‘most likely case’ [15]: Given their dedication both to broad and equal participation, and their commitment to making online participation processes work for all their members, if online participation should work anywhere, it is here. If it is successful, the research will provide insights as to why that is. If it is not, it will help understand why not, and identify pitfalls to avoid. Practical application of this research will include guidance for parties or organization who want to use online participation methods in a democratic context.

Contribution

This study will make several unique contributions:

- It provides an in-depth case study of the introduction of online participation in a political party. Though plenty of studies have looked at how parties engage citizens or local

members, very few have focused on internal processes, and none have done it at this scale.

- It combines four potential effects of the introduction of online participation methods – mobilization, reinforcement, replacement and non-use – which have not yet been considered in a single model. The study provide a dense evidence base to show all of these effects in action.
- A modest methodological contribution is made in a new type of measurement for participation preferences in surveys.

Acknowledgements

This research was funded by the EPSRC Centre for Doctoral Training in Web Science Innovation, EP/L016117/1. Funding for fieldwork was supplemented by the Green Party Germany. The author is not a member of the party, and the party had no influence on the analysis and conclusions drawn.

References

- [1] E. G. Frankland, “The Evolution of the Greens in Germany: From Amateurism to Professionalism,” in *Green Parties in Transition: The End of Grass-Roots Democracy?*, E. G. Frankland, P. Lucardie, and B. Rihoux, Eds. Farnham, Surrey: Ashgate, 2008, pp. 19–42.
- [2] R. Heinrich and M. Spitz, “Relevanz gewinnt - Online-Partizipation im Wahlkampf bei den Grünen!,” in *Internet und Partizipation*, Wiesbaden: Springer Fachmedien Wiesbaden, 2014, pp. 223–236.
- [3] M. Kellner, “Beteiligungspartei 2019,” 2015.
- [4] J. L. Jensen, “Political Participation Online: The Replacement and the Mobilisation Hypotheses Revisited,” *Scan. Polit. Stud.*, vol. 36, no. 4, pp. 347–364, Dec. 2013.
- [5] G. Vowe, “Digital Citizens und Schweigende Mehrheit: Wie verändert sich die politische Beteiligung der Bürger durch das Internet? Ergebnisse einer kommunikationswissenschaftlichen Langzeitstudie,” in *Internet und Partizipation*, no. Mpfv 2011, Wiesbaden: Springer Fachmedien Wiesbaden, 2014, pp. 25–52.
- [6] S. Halford and M. Savage, “Reconceptualizing Digital Social Inequality,” *Information, Commun. Soc.*, vol. 13, no. 7, pp. 937–955, Oct. 2010.
- [7] P. Norris, *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*. Cambridge: Cambridge University Press, 2001.
- [8] M. Margolis, D. Resnick, and C.-C. Tu, “Campaigning on the Internet - Parties and candidates on the World Wide Web in the 1996 primary season,” *Harvard Int. J. Press.*, vol. 2, no. 1, pp. 59–78, 1997.
- [9] S. Vissers and D. Stolle, “The Internet and new modes of political participation: online versus offline participation,” *Information, Commun. Soc.*, vol. 17, no. 8, pp. 937–955, 2014.
- [10] C. Lutz and C. P. Hoffmann, “The dark side of online participation: exploring non-, passive and negative participation,” *Information, Commun. Soc.*, vol. 20, no. 6, pp. 876–897, 2017.
- [11] A. Kerr and J. Waddington, “E-Communications: An Aspect of Union Renewal or Merely Doing Things Electronically?,” *Br. J. Ind. Relations*, vol. 52, no. 4, pp. 658–681, Dec. 2014.
- [12] K. Gerl, S. Marschall, and N. Wilker, “Does the Internet Encourage Political Participation? Use of an Online Platform by Members of a German Political Party.,” *Policy & Internet*, vol. 9999, no. Im Erscheinen, 2017.
- [13] G. Thuermer, S. Roth, K. O’Hara, and S. Staab, “Online participation in democratic processes: The case of the Green Party Germany,” 2017.
- [14] G. Thuermer, S. Roth, K. O’Hara, and S. Staab, “Everybody thinks online participation is great – for somebody else,” in *Proceedings of the 10th ACM Conference on Web Science*, 2018, pp. 1–10.
- [15] B. Flyvbjerg, “Five Misunderstandings About Case-Study Research,” *Qual. Inq.*, vol. 12, no. 2, pp. 219–245, 2006.

Tag Thunder: Towards new oral reading strategies for visually impaired people

Judith Jeyafreeda Andrew
University de Caen Normandie
Caen, France
judith-jeyafreeda.andrew@unicaen.fr

Stéphane Ferrari
University de Caen Normandie
Caen, France
stephane.ferrari@unicaen.fr

Fabrice Maurel
University de Caen Normandie
Caen, France
fabrice.maurel@unicaen.fr

ABSTRACT

The scope of the PhD lies within the Tag Thunder Project (TTP). It aims to allow accessibility of web page for the Visually Impaired People (VIP). TTP uses these guidelines to allow skimming and scanning of a web page for the visually impaired people. For this purpose, our platform produces a tag thunder, which is an oral transposition of the tag cloud concept. First, it retrieves representative key-terms from the web page which are then vocalized and spatialized by considering different factors to form the Tag Thunder (area of origin, typographic clues, position, number of occurrences...). The PhD is specifically concerned with two modules of the TTP - one, for breaking down the web page into a set of zones each consisting of coherent information both visually and thematically and the other for extracting representative words from each zone, words which deliver the general idea behind the zone.

KEYWORDS

Web Accessibility, Segmentation, Extraction, Tag Thunder

ACM Reference Format:

Judith Jeyafreeda Andrew, Stéphane Ferrari, and Fabrice Maurel. 2018. Tag Thunder: Towards new oral reading strategies for visually impaired people. In *Proceedings of 10th ACM Conference on Web Science (WebSci)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 INTRODUCTION

As the Internet develops, the web related applications have become one of the most significant applications of networks. A web page has to be accessible to everyone. In this paper, we define two partially unconscious/automated processes - skimming and scanning. Skimming is the process deeply involved to elaborate a global reading strategy; and scanning in the capacity to jump from a zone of interest to another. Tag thunder is a project that aims at allowing VIP to use and combine these human natural abilities into efficient strategies, as well as a person with sight.

The skimming and scanning process is based on several factors like layout, logical structure and typographic effects which are available in the visual environment. However, these features are very limited in the non-visual environment, as shown in Figure 1:



Figure 1: Sight vs. Blind perception of web page (Allowing non sequential vs. Imposing sequential reading strategies)

a same web page (on the left), made to be perceived by screen readers (on the right), has lost much structural information which enables easy, rapid and natural exploration of the document in a non sequential manner, thus making skimming and scanning a difficult, if not an impossible task.

[10] gives a detailed account on Tag Thunder process, which takes a web page as input, integrates in a single file all the visual information provided by several source, identifies a given number of zones and representative key terms for each of them which are vocalized and spatialized by considering different factors to form the Tag Thunder (area of origin, typographic clues, position, number of occurrences...), which allows the visually impaired to skim and scan a web page. Figure 2 shows the different modules of TTP. However, the PhD is targeted in the information retrieval phase (modules in pink), which are the web page segmentation and extraction of key ideas phase of Tag Thunder. Though the PhD will address both the topics, at present the work is concentrated on the task of segmentation.

2 OBJECTIVES

In order to enable skimming and scanning for VIP, the segmentation task in TTP will include the three features - visual, logical and semantic, of web pages. This will be done by taking advantage of the structure of the web page obtained using the HTML DOM.

The user should be able to interact with the system. We want to allow the user to make the choice of what they want to read, and to be able to zoom in and zoom out of the web page.

For the extraction tool, the objective is to select key words or propose a summary for each zone. The extraction tool will be worked on after the zoning of the web pages are achieved.

3 HYPOTHESIS

There are several different techniques for segmentation: visual approaches - DOM based: [3], [2], [9], [6], image based: [4] clustering

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

WebSci, May 2018, Amsterdam, Netherlands

© 2018 Copyright held by the owner/author(s).

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.

<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

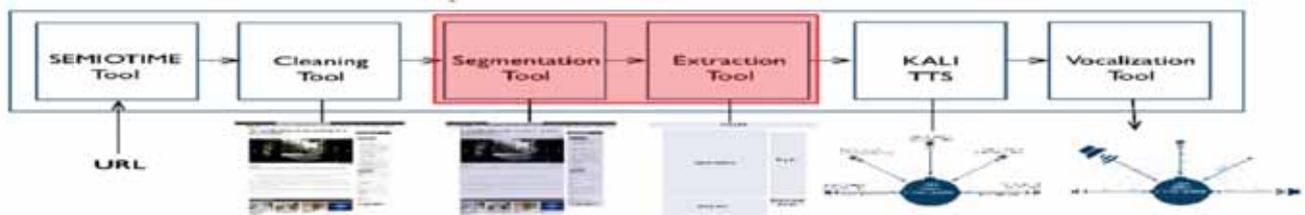


Figure 2: Architecture of Tag Thunder



Figure 3: Segmented page with key terms extracted

technique - [5], hybrid approaches - [11], [12]. However, it is noticed that there is no technique that considers all the three aspects (visual, layout semantic). Also, most techniques classify the web page as noisy and content zones, and tend to eliminate the noisy zones (citations, advertisements etc.). But, we want to allow the user to make the choice of what he/she wants to read. Also, we want a certain number of zones, such that we can use the cocktail party effect.

Thus, we aim at using Constraint based clustering for segmentation, thus allowing us to define constraints [7] based on the visual, logical and semantic features of the web elements. Some of them will be inspired by the Gestalt theory [8]. We also plan to include constraints based on the guidelines of web accessibility proposed by [1], wherever possible. The constraint based clustering requires a similarity measure, which should also take into account the same three features as well. The segmentation technique by constraint based clustering is expected to produce an output as depicted in the left of Figure 3.

We also plan to implement other techniques for segmentation and the hypothesis is that, certain techniques would work well on certain type of web pages depending on the structure of the web page. Therefore, ensemble learning will be implemented to find exact methods which suits a particular web page.

Although the extraction tool will be handled after the segmentation tool, the expected output from the extraction tool is as depicted in the right of Figure 3. Key terms are extracted from each zone that has been created by the segmentation tool. The user should also be able to interact with the tool by zooming in and out of any particular zone. While doing so, the Tag Thunder does the segmentation and extraction process again to present the user with a new set of key terms.

For evaluation, we plan to work on a measure to automatically verify and quantify some specific criteria, basically the same used to build the dissimilarity matrix for the constraint-based clustering.

However, there still remains the problem of weighting the criteria. They all do not have the same importance and probably can not be satisfied at the same time. Therefore, the issue is how they should be balanced

4 WORK PLAN

The PhD work for the first year will be concentrated on the segmentation phase. During the first three months of the PhD (October 2017 to December 2017) an intensive literature work was done to learn and compare the existing works. The months January 2018 to February 2018 was dedicated to identifying the atomic elements of a web page using the HTML DOM structure, in order to be considered for the segmentation process. The months March 2018 - July 2018 have been dedicated to conducting experiments with different criteria and similarity measures. Through the months August 2018 to December 2018 ensemble learning will be used to compare different techniques on different web pages.

REFERENCES

- [1] [n. d.]. WCAG Guidelines. <https://www.w3.org/WAI/intro/wcag>. ([n. d.]).
- [2] Ms S Aruljothi, Mrs S Sivaranjani, and S Sivakumari. [n. d.]. Web Page Segmentation for Small Screen Devices Using Tag Path Clustering Approach. ([n. d.]).
- [3] Deng Cai, Shipeng Yu, Ji-Rong Wen, and Wei-Ying Ma. 2003. *VIPS: a Vision-based Page Segmentation Algorithm*. Technical Report MSR-TR-2003-79. Microsoft Research.
- [4] Jiuxin Cao, Bo Mao, and Junzhou Luo. 2010. A segmentation method for web page analysis using shrinking and dividing. *International Journal of Parallel, Emergent and Distributed Systems* 25, 2 (2010), 93–104.
- [5] Freddy YY Choi. 2000. Advances in domain independent linear text segmentation. In *Proceedings of the 1st North American chapter of the Association for Computational Linguistics conference*. Association for Computational Linguistics, 26–33.
- [6] Juntang CHEN Cunhe LI, Juan DONG. 2010. Extraction of Informative Blocks from Web Pages Based on VIPS. (01 2010).
- [7] Khanh-Chuong Duong, Christel Vrain, et al. 2013. A declarative framework for constrained clustering. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 419–434.
- [8] Willis Davis Ellis. 1967. *A source book of Gestalt psychology*. Vol. 2. Psychology Press.
- [9] Xiao-Dong Gu, Jinlin Chen, Wei-Ying Ma, and Guo-Liang Chen. 2002. Visual based content understanding towards web adaptation. *AH* 2 (2002), 164–173.
- [10] Jean-Marc Lecarpentier, Elena Manishina, Fabrice Maurel, Stéphane Ferrari, and Gaël Dias. 2016. Tag Thunder: Web Page Skimming in Non Visual Environment Using Concurrent Speech. In *7th Workshop on Speech and Language Processing for Assistive Technologies (SLPAT 2016) associated to INTERSPEECH 2016*. San Francisco, United States. <https://hal-preprod.archives-ouvertes.fr/hal-01496711>
- [11] Cong Kinh Nguyen, Laurence Likforman-Sulem, Jean-Claude Moissinac, Claudie Faure, and Jérémy Lardon. 2012. Web document analysis based on visual segmentation and page rendering. In *Document Analysis Systems (DAS), 2012 10th IAPR International Workshop on*. IEEE, 354–358.
- [12] Premavathi M Rajdeeba B. 2014. segmenting web pages using correlation clustering and reducing noisy data using simple K means algorithm. *International Journal Of Innovative Research In Electrical, Electronics, Ins Trumentation And Control Engineering* (02 2014).

A Study on the Emergence of Online Bar Exam Training in China

LI XINGFANG, The University of Hong Kong, Hong Kong, u3508660@hku.hk

This paper examines how, and to what extent China's online bar exam training helps with the construction of high-quality professional legal education system in China. Respective characteristics and relations of online bar exam training and Chinese judicial examination will be explored. The rationales that behind such relation will be examined from the perspectives of changes of social environment, legal development and the trend of professionalism.

KEYWORDS

Online bar exam, legal education reform, trend of professionalism, internet technology – driven effect on legal professionals

Legal education is the education of individuals in the principle, practices and theory of law. In China, a big market of bar exam training exists. Unlike traditional legal education, bar exam training is particularly designed for helping individuals who want to pass Chinese bar exam. Chinese Bar exam is a test intended to determine whether a candidate is qualified to practice law in China. With the development of Internet and communication technology, we could find bar exam training through different channels in China nowadays, including massive open online course, small private online course, group chat, Internet broadcast, etc.

This paper examines the question as how, and to what extent China's online bar exam training helps with the construction of high-quality

professional legal education system in China. Starting from the general introduction of the development of (online) bar exam training in China, it analyses the respective characteristics and relations of online bar exam training and judicial examination. Several bar exam training agencies, such as *Houda*, *Reida* etc, will then be taken as examples to explore the rapid development of China's online bar exam training market, and how such training contributes to the current legal education reforms. Sample training agencies' operating mechanisms, self-created practice database, and target audience etc. will be examined in details.

It is argued in the paper that in order to truly implement China's legal education reforms and better regulate Chinese professional legal community, the reform of Chinese bar exam shall be deepened, with particular emphasis on fostering online bar exam training that facilitates the traditional legal education provided by law schools and other institutions. Online bar exam training, as a new type of professional legal education which bears the advantages of flexibility, low costs and time-saving, is suggested to be incorporated into the overall legal education system. Due to its unique technological, long-distance, and virtual community characteristics, both self-generated and formal regulations are needed to better manage online bar exam training in China.

Information Access for Development: Weather Information Access for Rural Ghana

Francis Dittoh
Vrije Universiteit Amsterdam
Amsterdam
The Netherlands
f.dittoh@vu.nl

Victor de Boer
Vrije Universiteit Amsterdam
Amsterdam
The Netherlands
v.de.boer@vu.nl

Hans Akkermans
Vrije Universiteit Amsterdam
Amsterdam
The Netherlands
hans.akkermans@akmc.nl

ABSTRACT

With very low internet penetration, over 64% of Africans have no internet access. In recent years, urban Africa has improved in terms of internet access with 3G and 4G available in some cities. Rural and suburban Africa however still has virtually no access to internet services. Issues of illiteracy, numerous local languages and the lack of infrastructure are the foremost barriers that keep these people unconnected although information needs such as market prices and weather forecasts might be relevant to these communities. These conditions are similar for most countries in sub-Saharan Africa. Information of this nature, available on the internet or elsewhere, can be accessed by proxy rather than directly. The widespread availability of mobile phones and radio access in rural areas suggests attention should be concentrated on these technologies because it is important to use what is readily available. This research will delve practically into how ICTs can provide rural communities with regular access to up-to-date information, from the world wide web, peers or organizations, which is relevant to their needs. The study will use an adapted living-labs approach which will involve all stakeholders, especially the end-users and institutions that deal with rural communities, in the development of various use-cases that are usable by the illiterate, accessible in their own language(s) and requiring inexpensive equipment for the system design and already existing and/or inexpensive equipment for the end-users. The research will also look at the various ways of creating technical sustainability, possibly with training and the development of communities of local developers. This paper focuses on the Weather Information Use-case

KEYWORDS

Web Access, literacy, development, ICT4D, Weather, Mobile, Radio, sub-Saharan Africa

1 INTRODUCTION

Information access is not readily available to everyone. Only about 35.2% of Africans have access to the internet [1]. This means a whopping 834,568,485 people in Africa have never accessed information on the Information Super Highway. Radio, TV and Mobile Telephony of course caters to the information needs of a large portion of this number, but many of the advantages of the Web are not realized by this large population. This research aims to answer the question of how ICTs can provide rural communities in Ghana (and other sub-Saharan Africa countries) with regular access to up-to-date information, from the world wide web, peers or organizations, which is relevant to their needs. The research will initially focus on Northern Ghana which covers 97,702Km² (40.96% of the nation's surface area) [1]. We select this area for our case study because its rural areas fit the targeted group, being an agriculture-production region and therefore having impact on the nation's food security and being more deprived of infrastructure, and with higher illiteracy.

2 RELATED WORK

2.1 ISKA

Ignitia Tropical Weather Forecasting¹ has a product known as ISKA, which runs in Ghana to provide weather forecasts through mobile phones. Users utilize a USSD code to subscribe and from then on receive weather forecasts via SMS. Needless to say, this, like many other existing ICT solutions, by way of design does not cater for the illiterate rural.

2.2 Smart Weather Station for Rural Agriculture using Meteorological Sensors and Solar Energy

Adoghe et al. (2017) look at access to short term weather forecasts for local communities in Africa. A cost-effective, solar-powered automated weather station was designed and implemented using

¹ <http://www.ignitia.se>

meteorological sensors, a microcontroller, a Liquid Crystal Display (LCD), and a GSM modem [3]. The resulting data can be read on the device and is also sent via SMS to user mobile phones. Once again, the issue of literacy remains a barrier.

3 INFORMATION ACCESS IN RURAL GHANA

3.1 Information Needs

Rural communities have information needs like any other region of the world. These needs are often related to their daily activities [4]. From various workshops and focus group discussions in Ghana, it is further confirmed that the information needs of rural folk center around daily activities which are mainly crop farming and livestock rearing; market access to offer produce for sale and buy farm inputs, the knowledge of prevailing market prices of produce and inputs, and obtaining weather forecasts; climate change affects their ability to follow previously normal trends.

3.2 Barriers

There are major barriers that have however made the above difficult to achieve. These include the *lack of infrastructure* which many ICT innovations would depend on, *low literacy rates* in rural areas, which means smart phones and computers, even if available, would seldomly be used and a myriad of *local languages* (Ghana has about 250), making the inability to understand online content difficult.

4 OPPORTUNITIES

There are however existing technologies that present opportunities for innovation in the area of information access.

4.1 Mobile Telephony

With a 131.9% Mobile Penetration [5] in Ghana, it stands to reason that voice telephony is widely used, more so in rural areas. This presents an opportunity of being able to reach everyone via voice-based technologies.

4.2 Radio

Radio has been in use all over Africa, and in rural areas, for decades and is widely used as the major source of information in rural areas. This presents an opportunity of being able to reach everyone via radio-based technologies.

4.3 Institutions

There are various institutions within Ghana that either already work with rural communities in various areas or are ready to do so. The flow of information to and from these rural communities is however often a challenge. Collaborations with these institutions is an opportunity to solve their challenges while creating sustainable systems vis-à-vis hosting and maintenance.

5 METHODOLOGY

The research will involve all stakeholders, especially the end-users and institutions that deal with rural communities in an adapted living-lab approach; involving the end user in the process of problem identification, technology design, implementation and evaluation is a more suitable approach for ICT4D Projects [6].

6 PROPOSED SYSTEM

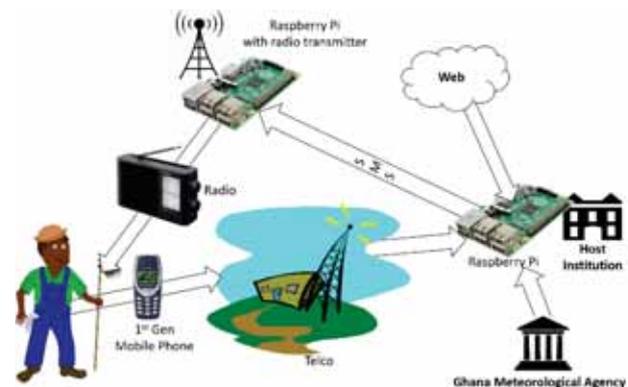


Figure 1: Conceptual Design of Weather Information System

The proposed system for this use-case will obtain weather forecast for a specific community (using GPS Coordinates) from the internet (open access weather information). Using concatenated pre-recorded files (local languages), it will create an audio fragment containing the obtained forecast. The audio fragment will then be made available on the system via a phone number, which users can call in to or a repeated FM broadcast over a short distance which users can tune in with any radio receiver.

REFERENCES

- [1] Internet World Stats, Retrieved April 2018 from <https://internetworldstats.com/stats1.htm>
- [2] Regions - Government of Ghana, Retrieved April 2018 from <http://www.ghana.gov.gh/index.php/about-ghana/regions>
- [3] Adoghe et al, 2017. Smart Weather Station for Rural Agriculture using Meteorological Sensors and Solar Energy, London, UK
- [4] S. Islam and Z. Ahmed, 2012, The information needs and information-seeking behaviour of rural dwellers: A review of research, IFLA Journal, Vol 38, Issue 2, pp. 137 - 147
- [5] National Communications Authority, 2016 Quarterly Statistical Bulletin on Communications in Ghana, Volume 1, Issue 2
- [6] Baelden and Audenhove, 2015, Participative ICT4D and living lab research: The case study of a mobile social media application in a rural Tanzanian University setting, Telematics and Informatics, Volume 32, Issue 4, November 2015, Pages 842-852