The 8th ACM Web Science Conference 2016

Robert Jäschke
L3S Research Center
Hannover, Germany

This article provides an overview of this year's ACM Web Science Conference (WebSci'16). It was located in Hannover, Germany, and organized by L3S Research Center and the Web Science Trust. WebSci'16 attracted more than 160 researchers from very different disciplines – ranging from computer science to anthropology. Celebrating 10 years of the Web Science research initiative, the conference featured six keynotes, three panels, nine paper sessions, and several side-events.

1. INTRODUCTION

The *Eighth International ACM Web Science Conference 2016 (WebSci'16)* was held from Sunday, May 22, to Wednesday, May 25, 2016 in Hannover, Germany. Celebrating two anniversaries, the conference featured an exceptional program with six keynotes, three panels, and several excellent side-events.

This year, 10 years of the Web Science research initiative were celebrated – an initiative which aims to study the people and technologies, applications, processes and practices that shape and are shaped by the World Wide Web. Web Science aims to draw together theories, methods and findings from across academic disciplines, and to collaborate with industry, business, government and civil society, to develop our knowledge and understanding of the Web: the largest socio-technical infrastructure in human history. The conference brought together researchers from multiple disciplines, like computer science, sociology, economics, information science, anthropology, and psychology. They gathered for four days to discuss about all aspects of the Web, from the study of information networks, social communities, and organizations to applications and policies.

WebSci'16 was part of the celebrations of the Leibniz Year 2016, commemorating the 300th anniversary of the death of Gottfried Wilhelm Leibniz (1646-1716). Leibniz was living and working in Hannover for forty years and is the eponym of the Leibniz University Hannover, which has borne the polymath's name since 2006 and thus celebrated its tenth anniversary.

The conference was organized by Leibniz University's L3S Research Center, one of the 19 world-class international research centers which form the Web Science Trust's WSTNet. The conference series by the Web Science Trust is following events in Athens, Raleigh, Koblenz, Evanston, Paris, Indiana, and Oxford.

This article provides a brief summary of WebSci'16 from the perspective of a participant who was also the Local Chair of the conference.

2. ORGANIZATION

WebSci'16 was chaired by the general chairs Dame *Wendy Hall* (University of Southhampton, UK) and *Wolfgang Nejdl* (L3S Research Center and Leibniz University Hannover, Germany). The program chairs were *Paolo Parigi* (Stanford University, USA) and *Steffen Staab* (University Koblenz-Landau, Germany and University of Southampton, UK)

At such an interdisciplinary conference as WebSci, the PC chairs have the tough job to find adequate reviewers, moderate between different disciplinary cultures, and balance the coverage of topics. *Steffen Staab* and *Paolo Parigi* did an excellent job. They were backed by 12 senior PC members with diverse disciplionary backgrounds who suggested reviewers and lead the discussions. This year we were very pleased to have received 70 full paper submissions, 48 short paper submissions, and 16 extended abstracts. Almost all submissions received more than 3 reviews, for a total of over 550 reviews from around 86 reviewers. Based on the reviews, discussions among the reviewers, and a meta-review phase by the senior PC members and the program chairs, 13 submissions were accepted as long talks, 23 as short talks, and 30 as posters For the first time at WebSci an award for the best reviewer was awarded. The happy winner was *Isabella Peters* from the University of Kiel, Germany, who was honored for her outstanding and helpful reviews by the PC chairs.

The conference started on Sunday, May 22, with workshops and tutorials. On Monday morning, it was opened by *Stefan Schostok*, the Major of Hannover, *Monika Sester*, the Vice President of Leibniz University Hannover, and the general chair *Wolfgang Nejdl*. From Monday to Wednesday the more than 160 participants could attend several keynotes, panels, paper sessions, an entrepreneurshop track, a hackathon, and a poster reception.

WebSci'16 purposely started two days after the *10th International AAAI Conference on Web and Social Media (ICWSM)* which took place at GESIS in Cologne, from May 17–20. The intention was to enable researchers interested in the topically related conferences to easily visit both of them, since Cologne and Hannover are separated by only two and a half hours of a convenient train ride. A post-conference comparison of the participants resulted in 20 joint participants, 12 of them from abroad, which is a good portion of the WebSci'16 participants and suggests that this could be a model for future editions.

3. KEYNOTES

A corner stone of the general chairs' goal to compile an attractive program for celebrating 10 years of the Web Science intiative was to invite top-class keynote speakers. With the line-up featured at WebSci'16 they clearly succeeded.

The conference was opened by *Ricardo Baeza-Yates*, former vice president of research at Yahoo! and currently affiliated with the Universitat Pompeu Fabra, Spain, and the Universidad de Chile. His keynote "Data and Algorithmic Bias in the Web" [Baeza-Yates 2016]showed both impressive and frightening examples of bias on the web. He emphatically reminded us that "the quality of any algorithm is bounded by the quality of the data it uses" and argued for better addressing bias in data.

On Tuesday *Jure Leskovec* (Assistant Professor of Computer Science at Stanford University and Chief Scientist at Pinterest) in his keynote "Antisocial Behavior in Online Communities" showed how trolling users in online communities can influence other users and how they are influenced. Supported by the finding that users which receive feedback are more likely to become engaged, even more with negative feedback, he ended his keynote with the request to "keep calm and don't feed the troll". The closing keynote of the day was by *Daniel Olmedilla* (Engineering Manager at Facebook) on "Applying Machine Learning to Ads Integrity at Facebook" [Olmedilla 2016]. Daniel provided insights into how Facebook leverages a combination of automated machine learning models and human computing to detect policy violating and low quality ads.

Wednesday, the last day of WebSci'16, was packed with three keynotes covering very diverse topics. *Daniel Miller* (Professor of Anthropology at University College London) started the day with his keynote "Why we Post – The Comparative Anthropology of Social Media" [Miller 2016]. He presented results from a large ethnographic study, in which nine anthropologists spent fifteen months in nine different communities in China, Brazil, Turkey, Chile, India, England, Italy and Trinidad. They found very diverse reasons and consequences on people's life, which they have summarized in open access books – one desribing each community. Daniel gave the WebSci'16 participants an insight from a perspective which is not yet so common at the conference but feedback showed that his keynote was most welcome and relevant for the community. The second keynote "Large-Scale Analysis of Dynamics of Choice Among Discrete Alternatives" [Tomkins 2016] by *Andrew Tomkins* (Engineering Director at Google Research) brought us back to the technical challenges that are driving research on the Web. Andrew presented insights on the problems recommender systems and similar algorithms face in situations where users have a variety of alternatives and where various factors influence their decision. The closing keynote was given by *Helen Margetts*, Director of the Oxford Internet Institute and Professor of Society and the Internet. In her keynote "The Data Science of Politics: How Social Media Shape Collective Action" [Margetts 2016] Helen discussed which influence social media has on political systems and how Web Science can help us to understand, explain and forecast political turbulences.

4. PANELS

Three panels provided opportunities to spark a dialogue and support interdisciplinary exchange. On Monday, the panel "Privacy and Internet Governance" chaired by *Nikolaus Forgó* from Leibniz University Hannover discussed the challenges for privacy induced by the pace of innovation [Forgó 2016]. Special guest was Max Schrems, who with his "Europe vs. Facebook" lawsuit led to the reconsideration of the Safe Harbour agreement between the US and EU. In the afternoon, *Steffen Staab* chaired the "10 Years of Web Science" panel in which *David De Roure Susan Halford*, *Anni Rowland-Campbell*, and (via video phone) *Jim Hendler* discussed what Web Science has achieved and what challenges it is currently facing [Hall 2016]. Finally, on Tuesday the panel "Computational Social Science: a *bricolage* of approaches" chaired by *Paolo Parigi* and with *Jeff Hancock*, *Claudio Cioffi-Revilla*, and *Davide Bennato* presented different approaches of this burgeoning field of research [Parigi 2016].

5. PAPER SESSIONS

The 33 presentations were grouped into the nine sessions *Behavior*, *Digital Rights and Public Access*, *Information Spreading and Engagement*, *Politics and the Web*, *Categorization and Predictions*, *Information Spreading*, *Information Gathering*, *Conceptualization*, and *Relationships and Entity*. All sessions were single-track, except for four sessions for short papers. This year we experimented with a new session model which was motivated by H. V. Jagadish's CACM article [Jagadish 2016]. The goal was to have more time for discussion and to identify connections between different research topics. At the beginning of each session, the session chairs introduced all speakers and papers and set them into the context of the session's topic. Then presentations of 15 minutes per paper (8 minutes for short papers) followed. Each session finished with a 30 minute panel which allowed the audience to get involved by asking questions and contributing to the discussion. Both the session chair and the presenters had read all papers and also prepared questions to spark the discussion. The reactions to this model were mixed. For a first try, it worked surprisingly well but there is clearly more experience necessary. One challenge is the heterogeneity of the topics at WebSci, which makes it difficult to find connections. After four presentations attendees also found it hard to remember the first talk and have questions to its speaker – a drawback also for those presenters, since they often did not receive feedback. We performed a survey among participants which will help next year's organizers. We will see whether this experiment can evolve into a successful WebSci session model.

Five papers were nominated for the Best Paper Award, the two winners are:

—Understanding Video-Ad Consumption on YouTube: A Measurement Study on User Behavior, Popularity, and Content Properties by *M.V. Siqueira de Arantes, F. Figueiredo, J. Almeida* [Arantes et al. 2016]

—Identity Assurance in the UK – technical implementation and legal implications under the eIDAS Regulation by *N. Tsakalakis, K. O'Hara, S. Stalla-Bourdillon* [Tsakalakis et al. 2016]

6. SIDE-EVENTS

Besides the traditional *tutorials* and *workshops*, WebSci'16 also featured two novel formats: a *hackathon* and an *entrepreneurship track*.

The workshop chairs *Ingmar Weber* (QCRI, Qatar) and *Bogdan State* (Facebook and Stanford University, USA) selected four *workshops* which took place on Sunday:

—Online Safety, Trust and Fraud Prevention, organized by *Marco Fisichella, Mario Elstner, Ricardo Kawase*, and *Tobias Knuth*

—Weaving Relations of Trust in Crowd Work, organized by *Cristina Sarasua, M. Six Silberman, Kristy Milland*, and *Gianluca Demartini*

—Natural Language Processing and Computational Social Sciences, organized by *David Bamman, A. Seza Doğruöz, Jacob Eisenstein, Dirk Hovy, David Jurgens, Brendan O'Connor, Alice Oh, Oren Tsur*, and *Svitlana Volkova*

—Data Driven Innovation on the Web, organized by *Gareth Beeston, Xin Wang*, and *Christopher Phethean*

Attending two workshops only, I can say that both had very interested and active attendees. In the "Weaving Relations of Trust in Crowd Work" workshop the keynote speaker *Kristina Holstein* from mobile.de showed how difficult it is to protect users of web marketplaces from fraud and which methods the companies leverage to improve trust in such platforms. A very diverse format encouraging interaction had the "Natural Language Processing and Computational Social Sciences" workshop with a keynote by *Claudia Wagner*, four paper presentations, and an afternoon roundtable discussion. After an introduction by the chairs, four groups formed to elaborate on the challenges that natural language processing faces in the context of computational social science. The suggestions and opinions on issues like reproducibility, ethics, or bias prevention were then exchanged in a plenary discussion.

In contrast to last years' editions, WebSci'16 featured five *tutorials* which were selected by tutorial chair *Claudia Wagner* (GESIS, Cologne, Germany):

—Utilising Online Qualitative Methods for Web Science, organized by *Laura Hyrjak, Amy Lynch, Sophie Parsons*, and *Lisa Sugiura*

—Likeology: Modeling, Predicting, and Aggregating Likes in Social Media organized by *Dongwon Lee*

—It's Getting Crowded! How to Use Crowdsourcing Effectively for Web Science Research organized by *Ujwal Gadiraju, Gianluca Demartiniy, Djellel Eddine Difallahz*, and *Michele Catast*

—Topic Model Tutorial organized by *Christoph Carl Kling, Lisa Posch, Arnim Bleier*, and *Laura Dietz*

—Community Detection: From Plain to Attributed Complex Networks organized by *Martin Atzmueller*

The novel *Entrepreneurship Track* was an ideal opportunity for startups to get in touch with international researchers and for researchers to see how ideas can successfully be implemented. The event was organized by *Susanne Oetzmann* and *Gabriele Hermann-Krotz* (both from L3S Research Center). After a keynote by *Reza Asghari*, who is Professor of Entrepreneurship at the Technische Universität Braunschweig, four different Web startups presented their experience on how to successfully transfer an idea into a commercial endeavour. The topics ranged from testing of new ideas (betatestr) and web usability (nurago) to web-based end-to-end encryption (Tutao) and metadata obfuscation (Qabel).

Another novel side-event which attracted a lot of attention was the *Alexandria and Archive-It Hackathon* organized by *Avishek Anand* and *Helge Holzmann* from L3S Research Center and *Jefferson Bailey* from the Internet Archive. After introductory sessions on Monday, four teams started to analyze Web collections ranging from 100 gigabytes to several terabytes in size and covering topics like the Occupy movement, the Fukushima nuclear disaster, or the Ebola outbreak. Using open source cluster management tools, libraries like ArchiveSpark and WarcBase, and their programming skills, the teams implemented amazing analysis platforms. Two successful teams received an award for the results they achieved during the three days: *Vinay Setty*, *Andreas Gerlach*, *Sarath Kondreddi*, *Natalia Boldyrev*, and *Paras Mehta* received the *Busy Beaver Award* for their *Spatio-temporal visualization of the Occupy movement*. They analyzed the IP addresses and locations of the

servers in the Occupy Movement collection and created an impressive visualization.[1] *Tarcisio Souza*, *Jaspreet Singh*, and *Fillipe Reis* received the *Outrageous Idea Award* for their work *Extracting symptoms for diseases from patient doctor accounts*.

7. SOCIAL EVENTS

The venue was situated at the shore of the Maschsee, an artificial lake in the city of Hannover. Participants enjoyed the large terrace which was a good place for coffee breaks, informal meetings, or just to relax. On Sunday evening I offered to guide participants through the city center and show them the nice and interesting spots of Hannover. I was overwhelmed by the large number of attendees and the positive feedback (despite my rather anecdotal knowledge of Hannover's past). The tour ended close to the restaurant district such that everybody had the chance to get some real German food.

Monday evening offered a densely packed (topically and spatially) and very lively poster reception. Participants could get some food and drinks to walk along the 30 posters which were presented and get in touch. A novel idea was the "hiking poster": the lightweight poster walls enabled the presenters to bring their poster to the audience instead of waiting for them. The voting participants honored this with the award for the best poster for *Ujwal Gadiraju, Patrick Siehndel*, and *Stefan Dietze* for their poster "Estimating Domain Specificity for Effective Crowdsourcing of Link Prediction and Schema Mapping".

The highlight social event was the conference dinner on Tuesday evening, located in the impressive New Town Hall of Hannover (which tourists often mistake for a palace). It featured a barbecue with an excellent selection of sausages and meat as well as delicious dips. At nine o'clock the happy winners of the best paper, best poster, and best reviewer awards were announced and prizes and certificates were commissioned.

8. CONCLUSION AND OUTLOOK

WebSci is a great interdisciplinary conference about all aspects of the Web. Still, interdisciplinarity is challenging for both (potential) participants and organizers, due to the differences among the disciplines. Issues like different publishing modes, options for financially supporting participants, fair reviewing, flexible options for presentation, and so forth, need be discussed and adapted. The steering committee is aware of these challenges and will surely find creative ways in continuing the success of WebSci.

During the conference, some attendees were very active on social media. Besides posting pictures from the social events, they extensively enriched the scientific program of the conference by tweeting remarkable insights, photos from the slides, or their personal comments. Clearly outstanding were two participants: *John S. Erickson* (@olyerickson) with his comprehensive tweets and also many humorous and helpful hints for speakers, participants, and organizers, and *Axel Bruns* (@snurb_dot_info) with his live blogging which covered many of the WebSci'16 talks[2] – a much enjoyed contribution.

Having successfully organized WebSci'16, I would like to thank the sponsors – Facebook

[1] `https://twitter.com/appelgriebsch/status/735149293317857281`
[2] `http://snurb.info/taxonomy/term/160`

and GESIS – and everybody who helped to make it a success, in particular our student volunteers, seven Master students in Computer Science. A big thanks goes also to my colleagues at L3S Research Center without whom the conference would not have been possible: *Angelika van Agen* (local arrangements chair), *Stefan Dietze* (finances chair), *Ujwal Gadiraju* (publicity chair), and *Miroslav Shaltev* (infrastructure chair).

> After the game is before the game!
> *Sepp Herberger*, German philosopher & coach of the national soccer team

WebSci'17 will be at Rensselaer Polytechnic Institute (RPI), Troy, New York, co-chaired by *Deborah L. McGuinness* (Tetherless World Senior Constellation Chair and Professor of Computer and Cognitive Science at RPI) and *Peter Fox* (Tetherless World Constellation Chair and Professor of Earth and Environmental Science, Computer Science and Cognitive Science at RPI). The PC co-chairs are *Katharina Kinder-Kurlanda* (Leader of the Data Linking & Data Security' team at GESIS, Bonn, Germany and adjunct lecturer at the Institute for Web Science and Technologie at the University of Koblenz-Landau) and *Paolo Boldi* (Professor at the Computer Science Department of the University of Milano, Italy).

REFERENCES

ARANTES, M., FIGUEIREDO, F., AND ALMEIDA, J. M. 2016. Understanding video-ad consumption on YouTube: A measurement study on user behavior, popularity, and content properties. See Hall et al. [2016], 25–34.

BAEZA-YATES, R. 2016. Data and algorithmic bias in the web. See Hall et al. [2016], 1–1.

FORGÓ, N. 2016. Privacy and internet governance. See Hall et al. [2016], 6–6.

HALL, W. 2016. 10 years of web science. See Hall et al. [2016], 7–7.

HALL, W., NEJDL, W., PARIGI, P., AND STAAB, S., Eds. 2016. *WebSci '16: Proceedings of the 8th ACM Conference on Web Science*. ACM, New York, NY, USA.

JAGADISH, H. V. 2016. Paper presentation at conferences: Time for a reset. *Commun. ACM 59,* 3 (Feb.), 38–39.

MARGETTS, H. 2016. Understanding political turbulence: The data science of politics. See Hall et al. [2016], 2–2.

MILLER, D. 2016. Why we post: The comparative anthropology of social media. See Hall et al. [2016], 3–3.

OLMEDILLA, D. 2016. Applying machine learning to ads integrity at Facebook. See Hall et al. [2016], 4–4.

PARIGI, P. 2016. Computational social science: A bricolage of approaches. See Hall et al. [2016], 8–8.

TOMKINS, A. 2016. Large-scale analytics of dynamics of choice among discrete alternatives. See Hall et al. [2016], 5–5.

TSAKALAKIS, N., O'HARA, K., AND STALLA-BOURDILLON, S. 2016. Identity assurance in the UK: Technical implementation and legal implications under the eIDAS regulation. See Hall et al. [2016], 55–65.

Robert Jäschke is senior lecturer at the Information School of the University of Sheffield, UK, and associated member of L3S Research Center, Hannover, Germany. Before he started his position in Sheffield in July 2016, Robert was assistant professor for computer science at Leibniz University Hannover (2012-2016), lecturer at the University of Kassel (2010-2012), and research fellow at the Knowledge & Data Engineering Group of the University of Kassel (2005-2010). He received his PhD in computer science from the University of Kassel for his research on collaborative tagging systems. Robert Jäschke is co-founder and manager of the web-based social bookmarking system BibSonomy (`http://www.bibsonomy.org/`). With several thousand active users, BibSonomy is one of the top three scientific literature sharing platforms in the Web 2.0. Robert can be contacted at jaeschke@L3S.de and his website address is `http://www.kbs.uni-hannover.de/~jaeschke/`.