COVER FEATURE GUEST EDITORS' INTRODUCTION

Web Science: Now More Than Ever

> James Hendler, Rensselaer Polytechnic Institute Wendy Hall, University of Southampton Noshir Contractor, Northwestern University

g=f[n]; ...

The Web has transformed our lives, and a cursory review of recent headlines shows both our dreams and nightmares writ large as a result. Is it possible to retain the benefits through algorithms and apps for romance, books, music, and fitness, without suffering the devastation wrought by the hacking of health records, elections, and bank accounts? This special issue explores the many facets of Web Science, the study of the Web's impact on our society and technology, and how we can use it to achieve our dreams without living our nightmares.

n a 2006 Science article, Web inventor Tim Berners-Lee and several others (including two of the guest editors of this special issue), addressed the need for an interdisciplinary field centered around studying both the social and technical aspects of the World Wide Web, and particularly around the interaction between those areas.¹ The authors argued that

...the scale, topology and power of decentralized information systems such as the Web also pose a unique set of social and public policy challenges. ... Transparency and control over the complex social and legal relationships behind this information is vital, but require a much more well-developed set of models and tools that can represent these relationships.

That article, and several others that followed,²⁻⁶ led to an emerging field now known as Web Science, which has been growing in the years since that first publication.

At the time Web Science was first

proposed, many of the things we have come to take for granted in Web use were in their infancy. The intervening years, however, have brought the societal impacts of the Web to the forefront. The Web's potential for mobilizing populations was cited as major factor in 2010's "Arab Spring,"⁷ and online activism has continued to increase significantly since then. This has led to the "hashtag" phenomenon of "counterpublic networks," in which those without access to major media platforms use social media to widely propagate their messages.⁸

Similarly, Web technologies have also been used to influence thinking. Indeed, as early as 2010, Web Science researchers showed how the Web could potentially affect major elections.⁹ Recent controversies have arisen over the use of Facebook data by Cambridge Analytica and the onslaught of influential advertising by Russian companies, which had not been identified as such at the time. We can no longer ignore the Web's utility in influencing the public's frame of mind as a growing factor in political discourse around the world. Even just a few years ago, who could have imagined the leader of a major country communicating his thoughts on national policies and preferences through daily Twitter blasts? In short, the Web's power to amplify messaging and communication is now understood—but will it be used to benefit humanity? Or will it bring out our worst features?

Beyond the political sphere, the Web has disrupted the future of work.¹⁰ The Web's early successes relied on voluntary efforts, most notably wikis, games with a purpose,¹¹ and citizen science.¹² However, it has also ushered in a new genre of crowdwork platforms that serve as a general purpose marketplace for paid work (such as Mechanical Turk, Upwork, Freelancer, ManPower), specialized marketplaces targeted at specific professions (such as TopCoder, uTest, 99Designs), or specialized activities such as crowdfunding (for example, Kickstarter and Kiva) and crowdsourcing design (such as Threadless). The Web, fueled by peer production networks,¹³ flash teams,¹⁴ flash

GUEST EDITORS' INTRODUCTION

organizations,¹⁵ and enterprise social media such as Slack,¹⁶ has enabled the creation of a sharing economy¹⁷ that coordinates peer-to-peer marketplaces for activities such as accommodation rentals (such as Airbnb), transportation (such as Lyft and Uber), domestic chores (TaskRabbit), delivery (Postmates), and many more. Together, these marketplaces have resulted in the dramatic rise of the "(fre)e-lance economy," which may bring unprecedented work opportunities to citizens around the globe, but also raises deep concerns about workers' welfare and rights.¹⁸

has gone behind the screens¹⁹ and is itself sitting on top of the infrastructure of the Internet, extending the ways in which all of the information can be accessed.

With changes in the ways people access the underlying Web, we need to better understand the interplays between the Web, its users, information providers and gatekeepers, and the larger society that is increasingly affected by it. The pursuit of Web Science increasingly requires exploring multiple avenues to better understand the growing and changing construct that is the World Wide Web.

IT IS ALSO THE CASE THAT AS THE WEB HAS BEEN CHANGING THE WORLD, THE WORLD OF TECHNOLOGY HAS BEEN CHANGING THE WEB.

It is also the case that as the Web has been changing the world, the world of technology has been changing the Web. Whereas in 2006 the browser was the primary Web interface, since then the Web has evolved from being a primary vehicle of communication to being the development platform on which other systems rely. Access to social media sites such as Facebook, Weibo, or Twitter is increasingly afforded through the use of mobile phones or other devices. Voice-controlled "virtual assistants," such as Siri, Alexa, and Google Home are increasingly penetrating markets and becoming an ever more common way for people to access information. In this paradigm, the Web increasingly becomes the invisible infrastructure, it

IN THIS ISSUE

This issue features a collection of papers exploring a variety of aspects of Web Science. We include papers that provide exemplars of the kinds of research needed to understand how Web technologies are impacting and being impacted by our lives both on- and offline. We also include papers exploring core attributes of the Web itself, including privacy, access, engagement, and globalization.

In "Understanding Users' Privacy Attitudes through Subjective and Objective Assessments: An Instagram Case Study," Kyungsik Han, Hyunggu Jung, Jin Yea Jang, and Dongwon Lee explore how users think about the privacy of their information on the Web-based Instagram platform. The authors studied how users' attitudes and behaviors were affected by information about "privacy leakage" in social media platforms. Based on this study, the authors discuss the implications for how Web platforms can better support users in managing their private information. The article is a good example of the Web Science "axiom" that good social science can inform the technological design of the Web.

In "Bringing Citizens and Policymakers Together Online: Imagining the Possibilities and Taking Stock of Privacy and Transparency Hazards," John Gastil and Sascha D. Meinrath reflect on the challenges of increasing civic engagement online. They examine what would be involved in integrating some of the best online tools for public engagement, while also exploring the challenges in doing so, given the high variance in quality and openness of the tools involved. This is a good example of studying the tradeoffs between the positive aspects of Web use (namely, how a system could boost the capacity and increase the public legitimacy of civic portals) and the associated challenges (including the related privacy and transparency issues).

In "How Do Organizations Publish Semantic Markup? Three Case Studies Using Public Schema.org Crawls," Daye Nam and Mayank Kejriwal present case studies of how different organizations bring semantic markup into their practices. They look closely at how schools, hospitals, and museums are increasingly using this tool, which embeds the machine-readable markup of the Semantic Web²⁰ onto Web pages via the schema.org initiative sponsored by major search engine companies.²¹ The paper not only presents some early, though intriguing, results, but is a good methodological example of Web Science in practice. In particular, the authors use information obtained from the Web Data Commons (http://webdatacommons .org), an open crawl that can be used for research studies such as this, to explore how the social and technical aspects of this emerging technology might affect users in the future.

In "Understanding Social Networks Using Transfer Learning," Jun Sun, Steffen Staab, and Jérôme Kunegis lay out a Web Science methodology used to more quickly understand user behaviors on emerging Web platforms. Traditionally, studies of any given Web platform must wait until enough data on the use of that system has been gathered and can be used by researchers. These authors show that transfer learning techniques can be used to see how information gathered on one or more current platforms can potentially help predict behaviors on other emerging platforms.

As the reach of the Web has grown, so too has its importance in connecting users into the global community, which creates a variety of problems for those communities in which connectivity is rare or nonexistent. In "Now You See It, Now You Don't: Digital Connectivity in Marginalized Communities," Carleen F. Maitland explores what happens when users are unable to get reliable access to the Web. in this case Native American communities in the US, as well as refugee communities in Rwanda and Jordan. The author studies these three groups and provides sociotechnical analyses of their Web usage. Maitland also explores the causes, systematic and otherwise, of problems in these communities, thus providing novel insights on aspects of the "digital divide," which have not previously received sufficient attention. The article also provides suggested solutions, both policy-based and technical, for those providing services to these communities. The author thus integrates the three pillars of Web science—social, policy, and technological aspects—for analysis, and proposes strategies to improve Web use within these communities.

In "Language Service Infrastructure on the Web: The Language Grid," Toru Ishida, Yohei Murakami, Donghui Lin, Takao Nakaguchi, and Masn summary, these articles comprise a snapshot of work going on in modern Web Science research and in the many disciplines involved. They explore policy implications and social impacts of the Web from a number of perspectives, ranging from qualitative social studies to quantitative mathematical analyses. These articles reflect the wide range of approaches needed to study and truly understand the nature of the Web and its impact on us all.

The major research challenges in achieving privacy, security, trust, and personal data protection in an open, interconnected platform are a major

THESE ARTICLES EXPLORE POLICY IMPLICATIONS AND SOCIAL IMPACTS OF THE WEB FROM A NUMBER OF PERSPECTIVES.

ayuki Otani provide less a study of Web use per se than a description of how systems engineered for the Web are being used to enhance interactions among groups of users who do not speak a common language. For more than a decade, the Language Grid project has been developing collaborative tools to support multilingual communities. This article reviews their work and describes how they have created an institutional structure for sharing language services between a large and growing community of providers and consumers. The architecture described in the article now supports 183 groups from 24 countries that share multilanguage-based services across distributed service grids.

focus for Web Science research. However, it is also clear that we are moving into an era in which our lives will be increasingly organized with the help of Web-based agents acting both as physical and thought partners. These agents will be powered by advances in artificial intelligence (AI) technologies, coupled with data gathered through decentralized networks of billions of sensors from the Internet of Things. Indeed these developments will not only impact the future of work but also the future of the workforce.²²

Web Science is just beginning to develop the research methodology to help us understand and enable these new networks of networks.²³ Understanding these changes requires an

GUEST EDITORS' INTRODUCTION

ABOUT THE AUTHORS

JAMES HENDLER is the Tetherless World Professor of Computer, Web, and Cognitive Sciences, and the director of the Institute for Data Exploration and Applications at Rensselaer Polytechnic Institute. His research interests include Web Science; applications of the Semantic Web; and artificial intelligence applications in science, engineering, and healthcare. He received his PhD in computer science from Brown University. He is a Fellow of IEEE, the AAAI, AAAS, ACM, and BCS. Contact him at hendler@cs.rpi.edu.

WENDY HALL is a professor of computer science and executive director of the Web Science Institute at the University of Southampton in the UK. Her research interests include multimedia, hypermedia, Web technologies, and Web Science. She received her PhD in pure mathematics at the University of Southampton. She is a Dame Commander of the Order of the British Empire, a Fellow of the Royal Society, the Royal Academy of Engineering, and the ACM, and a distinguished Fellow of the BCS. Contact her at wh@ecs.soton.ac.uk

NOSHIR CONTRACTOR is the Jane S. & William J. White Professor of Behavioral Sciences at Northwestern University. His research interests include understanding and enabling novel forms of organizing by leveraging and advancing network science, computational social science, and Web Science. He received his PhD in communication from the Annenberg School of Communication at the University of Southern California. He is a Fellow of the International Communication Association, a Distinguished Scholar of the National Communication Association, and a Distinguished Alumnus from the Indian Institute of Technology, Madras, where he received his undergraduate degree in Electrical Engineering. Contact him at nosh@northwestern.edu.

interdisciplinary approach. This field must simultaneously come to grips with the cyber-physical and AI worlds that our society is creating and anticipate their effects on our future—that is, how society co-creates, and is co-created by, the interaction between networks of people using highly inter-connected machines and emerging technologies.²⁴ We are only just beginning that journey.

REFERENCES

- T. Berners-Lee, et al., "Creating a Science of the Web," *Science*, vol. 313, no. 5788, 2006, pp. 769–771.
- N. Shadbolt and T. Berners-Lee, "Web Science Emerges," Scientific American, vol. 299, no. 4, pp. 76–81.
- K. O'Hara, et al, "Web Science: Understanding the Emergence of Macro-Level Features on the World Wide Web. Foundations and Trends in

Web Science, vol. 4, nos. 2–3, 2013, pp. 103–267.

- J. Hendler and W. Hall, "Science of the World Wide Web," *Science*, vol. 354, no. 6313, 2016, pp. 703–704.
- T. Tiropanis et al., "Network Science, Web Science, and Internet Science," Commun. ACM, vol. 58, no. 8, 2015, pp. 76–82; http://doi .org/10.1145/2699416.
- 6. T. Tiropanis et al., "Web Science Observatory," *IEEE Intelligent Systems*, vol. 28, no. 2, 2013, pp. 100–104.
- G. Lotan et al., "The Arab Spring, the Revolutions Were Tweeted: Information Flows during the 2011 Tunisian and Egyptian revolutions," Int'l J. Commun., vol. 5, 2011, p. 31.
- S.J. Jackson and B. Foucault Welles, "#Ferguson Is Everywhere: Initiators in Emerging Counterpublic Networks," Information Commun. Society, vol. 19, no. 3, 2016, pp. 397–418.
- P. Metaxas and E. Mustafaraj, "From Obscurity to Prominence in Minutes: Political Speech and Real-Time Search," Proc. Web Science Conf. (Web-Sci 10), 2010; www.researchgate .net/publication/228647699_From _Obscurity_to_Prominence_in _Minutes_Political_Speech_and _Real-Time_Search.
- A. Kittur et al., "The Future of Crowd Work," Proc. ACM Conf. Computer Supported Cooperative Work, 2013, pp. 1301–1318.
- L. Von Ahn and L. Dabbish, "Labeling Images with a Computer Game," Proc. SIGCHI Conf. Human Factors in Computing Systems, 2004, pp. 319–326.
- J. Raddick et al., "Galaxy Zoo: Motivations of Citizen Scientists," Astronomy Education Review, vol. 12, no. 1, 2013, pp. p010106-1–010106-27.
- 13. Y. Benkler, A. Shaw, and B.M. Hill, Peer Production: A Form of Collective Intelligence, MIT Press, Cambridge,

MA, 2015, pp. 175-204.

- D. Retelny et al., "Expert Crowdsourcing with Flash Teams," Proc. 27th Annual ACM Symp. User Interface Software and Technology, 2014, pp. 75–85.
- M.A. Valentine et al., "Flash Organizations: Crowdsourcing Complex Work by Structuring Crowds as Organizations," Proc. 2017 CHI Conf. Human Factors in Computing Systems (CHI 17), 2017, pp. 3523–3537.
- P.M. Leonardi, M. Huysman, and C. Steinfield, "Enterprise Social Media: Definition, History, and Prospects for the Study of Social Technologies in Organizations," J. Computer-Mediated Commun., vol. 19, no. 1, 2013, pp. 1–19.
- 17. A. Sundararajan, The Sharing Economy: The End of Employment and the

Rise of Crowd-Based Capitalism, MIT Press, 2016.

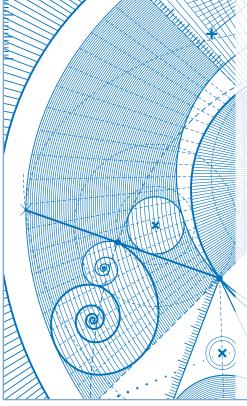
- S. Hill, Raw Deal: How the 'Uber Economy' and Runaway Capitalism Are Screwing American Workers, St. Martin's Press, 2015.
- W. Hall, "The Ever-Evolving Web: The Power of Networks," Int'l J. Commun., vol. 5, 2011, p. 14.
- T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," Scientific American, vol. 284, no. 5, 2001, pp. 34–43.
- R.V. Guha, D. Brickley, and S. Macbeth, "Schema.org: Evolution of Structured Data on the Web," *Commun. ACM*, vol. 59, no. 2, 2016, pp. 44–51.
- 22. National Academies of Sciences, Engineering, and Medicine,

Information Technology and the US Workforce: Where Are We and Where Do We Go from Here?, National Academies Press, 2017.

- 23. R. Ackland, Web Social Science: Concepts, Data and Tools for Social Scientists in the Digital Age, Sage, 2013.
- 24. J. Hendler and A. Mulvehill, Social Machines: the Coming Collision of Artificial Intelligence, Social Networks and Humanity, Apress, 2016.

DISCLAIMER

All three co-editors of this special issue are directors of the Web Science Trust (www.webscience.org).



Annals

of the History of Computing

From the analytical engine to the supercomputer, from Pascal to von Neumann, from punched cards to CD-ROMs—IEEE Annals of the History of Computing covers the breadth of computer history. The quarterly publication is an active center for the collection and dissemination of information on historical projects and organizations, oral history activities, and international conferences.

www.computer.org/annals