



Web Science Trust Observatory Project

**Wendy Hall
WEB SCIENCE TRUST**

**03/15/2018
Final Report**

DISTRIBUTION A: Distribution approved for public release.

**Air Force Research Laboratory
AF Office Of Scientific Research (AFOSR)/ IOE
Arlington, Virginia 22203
Air Force Materiel Command**

REPORT DOCUMENTATION PAGE			<i>Form Approved</i> <i>OMB No. 0704-0188</i>		
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Executive Services, Directorate (0704-0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.</p>					
1. REPORT DATE (DD-MM-YYYY) 02-08-2018		2. REPORT TYPE Final		3. DATES COVERED (From - To) 14 Mar 2018 to 14 Dec 2017	
4. TITLE AND SUBTITLE Web Science Trust Observatory Project			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER FA9550-15-1-0020		
			5c. PROGRAM ELEMENT NUMBER 61102F		
6. AUTHOR(S) Wendy Hall			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) WEB SCIENCE TRUST 40 Bank Street LONDON, E14 5DS GB			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) EOARD Unit 4515 APO AE 09421-4515			10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/AFOSR IOE		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-AFOSR-UK-TR-2018-0028		
12. DISTRIBUTION/AVAILABILITY STATEMENT A DISTRIBUTION UNLIMITED: PB Public Release					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The PI and her team were successful at initial deployment of technical, procedural, and policy elements toward creating a globally distributed and federated mechanism for sharing data and tools underpinning analysis of the Web, a so-called 'Web Observatory.' Her team's engagement with the community at a practical/hands-on level has been matched by internal work to establish key standards and tools behind the scenes. The Web Observatory serves not only as an object of study in its own right but as a tool to study the broader forces shaping the Web - as the Web, in turn, shapes behaviour of its users. Complete details are in the attached report. Also see www.webscience.org/web-observatory/ .					
15. SUBJECT TERMS Web Observatory Project, EOARD					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT SAR	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON NGUYEN, TRISTAN
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (include area code) 703-696-7796

Web Science Trust Observatory Project

Final PI Report 2017

Prof. Dame Wendy Hall, Web Science Trust



Dr. Ian Brown, Web Science Trust

USAF Award No. FA9550-15-1-0020

TABLE OF CONTENTS

SUMMARY	1
Acknowledgements	1
Management summary	2
Observing the Web	5
<u>A</u> Web observatory vs <u>the</u> web observatory	6
METHODS, ASSUMPTIONS, PROCEDURES	8
Method	8
Assumptions	8
Procedures	8
Deliverables by PHASE	9
RESULTS & DISCUSSION	14
Overview of outputs	14
Impact analysis	14
CONCLUSIONS & NEXT STEPS	17
Next Steps	17
REFERENCES	18
Bibliography	18

SUMMARY

ACKNOWLEDGEMENTS

The Web Science Trust would like to thank the Air Force Office of Scientific Research for supporting the work of the Web Science Trust and, in particular, for your continued support around the Web Observatory project.

On behalf of the Board of Trustees of the Web Science Trust:-

Prof. Dame Wendy Hall CEO, Web Science Trust

The authors would like to recognise the contributions, feedback and support delivered by the WO team, both at Southampton (led by Dr Thanassis Tiropanis) and also by the wider group of WSTNet members and Trust supporters internationally.

Prof. Dame Wendy Hall CEO, Web Science Trust

Dr. Ian Brown, University of Southampton

MANAGEMENT SUMMARY

This report covers deliverables/impact for the WO project in three main areas:

- *Refinement of the WO software infrastructure and tools*
- *Public engagement around WO at academic and public-facing events*
- *New published research into the underlying structure, context and application of WO.*

Our broad journey is from an early WO concept at the University of Southampton to a set of live international installations. The creation of standardised templates for Web Observatories based on our original SUWO proof of concept system for sharing data sets and analytics offers improved access and interoperability between research groups in academia, business and government.

Refined searching/sharing tools have enabled participants to discover Observatories and data sets in real-time using industry standard micro-notations such as schema.org.

A series of documentary and video resources have been developed to present the WO concept and large numbers of researchers/students have engaged in WO workshops, hackathons, conference workshops and summer school events internationally.

Our deliverables span three phases (roughly corresponding to 2015, 2016 and 2017)

Phase I (2015)

- *Development of a generic WO template from the Southampton design*
- *Establishing a micro data (schema.org) proposed standard for Web Observatory discovery*

- *Deployment of real-time crawling tools to report on the availability of WO systems, datasets and tools*
- *Authoring tutorial material for the template*
- *Installation of several working Web Observatory systems with partner universities*
- *Running a series of Web Observatory Workshops (WOW) and (BWOW) as well as a series of Web Observatory-focussed Datathons to engage with the user community.*

Phase II (2016)

- *Release of the new WO portal and landing pages*
- *Focus on complementary elements to WO adoption including licensing/legal issues, data discovery and the management of innovation adoption*
- *Expansion of the Web Science Network of Laboratories (WSTNet)*
- *Collaboration/engagement with WUN (Worldwide Universities Network) to deploy WO to all members for research data transparency and access plus scoping four demonstrator applications*
- *Running a series of Web Observatory and social machine workshops with the user community*
- *Several new publications bringing theory into practice*
- *A global media event (WebSci@10) featuring WO and the work of WST/WSTNet*
- *The creation of a Web TV channel resource*

Phase III (2017)

- *Additional WO installations through the WUN network*
- *Additional WO events (WebSci 17, WWW17)*
- *Featured WO article in Ten Years of Web Science*
- *Completion of the first PhD research project dedicated to WO*

- *Integration of the WO concept into broader themes such as the Internet of Things (IoT) and Artificial Intelligence (AI) and the Ethics of Big/Broad Data on the Web*
- *Delivery of WO learning material incl.*
 - *FutureLearn WebScience MOOC*
 - *FutureLearn Digital Marketing MOOC*
- *Ongoing/stable WO research streams internationally and additional research publications*
- *Delivery of final demonstrator WO back-end Observatory system and user portals (WebObservatory.org and WebObservatory.io)*

Introduction

OBSERVING THE WEB

Over the past 25 years the Web has grown to be a truly “world-wide” phenomenon meeting the ambition and, some might say, the arrogance of its name. As we see virtually all aspects of modern society either migrated or migrating to web-mediated alternatives we are faced with the challenge of answering a new set of questions about life on the Web beyond eCommerce. How can we understand what it means to govern, teach, commit crime, learn, play, research and profit on the Web? Where are the boundaries between transparency and privacy? Where does justified analysis become unjustified surveillance and how can we derive social benefits without giving up all the personal rights and privileges that underpin a free and open Web?

The capture of both structured/unstructured information on the Web (and about the Web and its users) forms the raw material for an interdisciplinary science of the Web beyond pure network science or big data analytics alone and brings together perspectives from maths, sociology, law, medicine, history, education, politics and beyond. This discipline, like other branches of science, needs tools as least as complex/powerful as the phenomena that are to be studied and the Web Science community has proposed that a virtual observatory similar to the virtual astronomical observatory (IVOA) would form the basis of such a tool.

A WEB OBSERVATORY VS THE WEB OBSERVATORY

The best analogy for (a) vs (the) Web Observatory is that of the Web itself – whilst there are many physical *web servers*, none of these servers is, itself, *the Web* – only part of the Web and more specifically part of the net effect of the interaction between all those web servers, their data/services and the people browsing them.

	A WO	THE WO (W³O)
Is a physical system	Yes	No
Is owned/operated by an authority	Yes	No
Draws data from more than one source	Maybe	Yes ¹
Contains Open Data	Maybe	Yes ²
Contains private data	Maybe	No ³

¹ Without more than one source it would be a reflection of a single WO

² Only the data that is open/shared will be visible on W³O

³ By extension although private data may exist on any/all WOs no private data would be available on the W³O

Whilst Web Observatories may be built only for this purpose the concept is intentionally more inclusive:

- *Systems may be explicitly designed and built to be Web Observatories and will probably use the term WO*

They may also have other names and other previous/parallel purposes and still contribute to the wider web of observatories, i.e. THE Web Observatory.

METHODS, ASSUMPTIONS, PROCEDURES

METHOD

The nature of the W³O project comprises an action research piece around the evolution and deployment of technical, procedural and policy elements of a globally distributed and federated mechanism for the sharing of data and tools underpinning analysis of the Web, its structure, its content and its users.

The key method is initially establishing standardised formats and protocols for the discovery and sharing of data and metadata between systems and then to encourage a network effect between these individual WO systems through demonstration, information sharing and engaging, to build a community of practice.

ASSUMPTIONS

- *There is NO assumption that a WO must be "branded" as such to participate – any system conforming to the minimal technical standards may participate*
- *There is NO assumption that all data or tools are hosted by the WO (they may instead be linked)*
- *There is NO assumption that all data on a WO must be open (though this clearly limits its shareability)*
- *There is NO assumption that all data/tools must be visible on a WO – though this may eliminate its contribution to the proposed network effects*
- *There is NO assumption fees/charges may not be levied for premium services or data*

PROCEDURES

There are no specific procedures documented for this project.

DELIVERABLES BY PHASE

Phase	Delivered	Description	Location/version /date
I	Technical Standard	Schema.org discovery	v1.0 (Current)
I		WO Template	V1.0 (Deprecated)
I		WO Crawler	V1.0 (Deprecated)
I		Discovery API	V1.0 (Deprecated)
I	Software Deployments	WO	Bangalore, India
I		WO	Seoul, S. Korea
I		WO	Perth, Australia
I		WO	Jakarta, Indonesia
I	Community Engagement	WO Event	Bangalore, India
I		WO Event	Seoul, S. Korea
I		WO Event	Perth, Australia
I		Datathon	Southampton, UK

I	Research Publications	"Building a Real-Time Web Observatory," R. Tinati, X. Wang, T. Tiropanis and W. Hall <i>(University of Southampton)</i>	IEEE Internet Computing, vol. 19, no. 6, pp. 36-45, Nov.-Dec. 2015.
II	Software Templates	WO template	v2 (updated)
II		WO crawler	v2 (updated)
II		Discovery API	v2 (updated)
II		WO Portal	v2 (updated)
II	Community Engagement	WO Event	WWW Workshop, Canada Apr 2016
II		WO Event	Data Ethics Workshop London 2016
II		WO Event	WebSci Workshop, Germany May 2016
II		WO Event	Global WebSci@10 Nov 2016
II		WO Comms	WO explainer Video
II		WO Comms	WO TV channel
II	WO Partnership	WUN WO project	http://wun.ac.uk/wun/research/view/web-observatory-project
II	Publications	Facilitating data-driven innovation using VOICE observatory infrastructure Eugene Siow, Xin Wang, Thanassis Tiropanis	May 2016 DDI '16: Proceedings of the Workshop on Data-Driven Innovation on the Web

		<i>(University of Southampton)</i>	Publisher: ACM
II		<p>Observlets: Empowering Analytical Observations on Web Observatory</p> <p>Aastha Madaan, Thanassis Tiropanis, Srinath Srinivasa, Wendy Hall <i>(University of Southampton)</i></p>	<p>April 2016 WWW '16 Companion: Proceedings of the 25th International Conference Companion on World Wide Web</p> <p>Publisher: International World Wide Web Conferences Steering Committee</p>
II		<p>Ethical and legal support for innovation on web observatories</p> <p>Caroline Wilson, Thanassis Tiropanis, Anni Rowland-Campbell, Leanne Fry</p>	<p>May 2016 DDI '16: Proceedings of the Workshop on Data-Driven Innovation on the Web</p> <p>Publisher: ACM</p>
III	Technical Standard	Schema.org discovery	v2 - updated
III		WO Template	v2 - updated
III		WO Crawler	v2 - updated
III		Discovery API	v2 - updated
		Porter Proxy	v1 - initial
III	Software Deployments	WO	MVP deployment webobservatory.io
III	Community Engagement	WO Event	WUN Workshop: Wellness Data for

			Healthy Societies, 2017
III		WO Conference	Brave Conversations. Canberra, Australia 2017
III		WO Engagement	web observatory.org
III		WO Proposal	The Role of Big Data in addressing the determinates o NC diseases in the ageing era
III	Research Publications	The DNA of Web Observatories Ian Brown	Mar 2017 Doctoral Thesis, University of Southampton
		Simon Price (<i>University of Bristol</i>), Wendy Hall Graeme Earl, Thanassis Tiropanis Ramine Tinati, Xin Wang Eleonora Gandolfi (<i>University of Southampton</i>), Jane Gatewood (<i>University of Rochester</i>) Richard Boateng (<i>University of Ghana</i>) David Denemark (<i>The University of Western Australia</i>), Alexander Groflin (<i>University of Basel</i>) Brian Loader (<i>University of York</i>) Maxine Schmidt, Marilyn Billings (<i>University of Massachusetts, Amherst</i>) Gerasimos Spanakis	Worldwide Universities Network (WUN) Web Observatory: Applying Lessons from the Web to Transform the Research Data Ecosystem. In R. Barrett, & R. Cummings (Eds.), WWW '17 Companion: Proceedings of the 26th International Conference on World Wide Web Companion (pp. 1665-1667). Perth, Australia.

		<p>(<i>Universiteit Maastricht</i>), Hussein Suleman (<i>University of Cape Town</i>), Kelvin Tsoi (<i>The Chinese University of Hong Kong</i>) Bridgette Wessels (<i>Newcastle University</i>) Jie Xu, Mark Birkin (<i>University of Leeds</i>)</p>	
III		Ramine Tinati, Aastha Madaan, Wendy Hall (<i>University of Southampton</i>)	The Role of Crowdsourcing in the Emerging Internet-Of-Things (WWW Conference 2017)
III		Xin Wang, Aastha Madaan, Eugene Siow, Thanassis Tiropanis (<i>University of Southampton</i>)	Sharing Databases on the Web with Porter Proxy (WWW Conference 2017)
III		Kashyap Popat, Subhabrata Mukherjee Jannik Strötgen, Gerhard Weikum (<i>Max Planck Institute for Informatics</i>)	Where the Truth Lies: Explaining the Credibility of Emerging Claims on the Web and Social Media (WWW Conference 2017)
III		Ramine Tinati, Aastha Madaan, Wendy Hall (<i>University of Southampton</i>)	InstaCan: Examining Deleted Content on Instagram WebSci '17
III		Jessica Ogden, Susan Halford, Leslie Carr (<i>University of Southampton</i>)	Observing Web Archives: The Case for an Ethnographic Study of Web Archiving WebSci '17

RESULTS & DISCUSSION

The results of the project so far are broken in to several areas including technical standards, software templates, software deployments, community engagement and impact analysis.

OVERVIEW OF OUTPUTS

IMPACT ANALYSIS

We list a number of references below in the growing body of research literature around observatories, data mining, big/broad data and note also the increased focus by government on the process of evidence-based policy creation and management.

The UK government has recently invested \$60m in a dedicated "Turing Institute" to promote data science across several UK universities and we see this trend increasing for the foreseeable future.

We have seen broad interest in Observatories across academia, business and government with the scale and diversity of data-oriented challenges growing. We anticipate more funding will be required to meet these challenges.

Several WOs have been created to form the core of a nascent W³O and we have ensured an inclusive approach to this making the discovery and display of datasets and tools independent of the underlying technologies.

There is a diversity of intent and perceived value models across the various groups engaging in this idea of observatories and more work is needed to compare/contrast the scope of ambitions in order to more easily construct an environment/ecosystem which fosters collaboration as well as simple interoperability at a technical level.

Given the objective to create network effects as we engage more systems and sources into the W³O eco-system, communication, training and practical tools for users like license and software templates are critical to support the adoption process. New research has been produced around innovation, licensing, simpler analytics, longer term work on WO and how its structure/processes relate to non-WO systems. This joins a growing body of research literature around observatories, data mining, big/broad data and note also the increased focus by government on the process of evidence-based policy creation and management.

The WO project with the Worldwide Universities Network (WUN) brings the issue of research data access and re-usability into focus and the WUN partners are working on four “vertical” WO demonstrators around four SIGs:

- **Massive Open Online Courses (MOOCS):** WUN members will be able to share resources about their MOOCS including datasets containing user interactions and communications, and visualisation tools to help understand how students participate.
- **Disaster Management:** As a prominent topic, the Web Observatory will provide a platform to share datasets and toolkits related to research within urban and natural disasters.
- **Government as a Social Machine:** the WUN network offers a fantastic opportunity to examine the Web Observatory in the context of Government as a Social Machine, where cross-institute, cross-Government analysis can be achieved by accessing an interconnected network of Observatories and datasets.
- **Young Citizens:** The WUN Networking Young Citizens interdisciplinary group will use the Web Observatory infrastructure to extend current research on a global scale including areas such as: India, East Asia and Africa.

Whilst its broader goal is to deploy WO templates to all WUN members for discovery, transparency and re-use of live (vs archived) research data assets, further underpinning the idea of WO as a live *social machine* for interdisciplinary research rather than as another static repository.

Another growing area of interest is on Trust in the provenance/quality of data with public outcry and government concern around “fake news” and the alleged manipulation of public opinion via social media distorting high profile democratic processes including the UK “Brexit” vote and US presidential election.

For WOs to achieve their potential they must offer four key forms of transparency supporting:

- *ease of access/use (functional transparency) especially for non-technical users*
- *clarity around provenance/quality information around underlying data (source transparency)*
- *clarity around conclusions/actions that follow from creating synthetic data/analytics (algorithmic transparency)*
- *accountability/permissions for producers and consumers of WO data and services (liability transparency)*

As the volume of data grows strongly and vastly exceeds our ability to engage with it, judge it and interpret it directly, we are left with a situation where many of the decisions made on the growing volumes of data will be made by algorithms rather than people and as the complexity of data and the realisation that decisions that are attempted will require more intelligence, nuance and flexibility. Thus the topic of Artificial Intelligence (AI) looms large around the processing of Web data and the ethics that underpin any “unattended” decisions.

Thus the Web Observatory serves not only as an object of study in its own right but as a vital tool to study the broader forces shaping the Web - as the Web, in turn, shapes our own behaviour.

CONCLUSIONS & NEXT STEPS

Our deployments of both hardware and software systems have been successful and our engagement with the community at a practical/hands-on level has been matched by internal work to establish key standards and tools behind the scenes. We look forward to new nodes, enhanced features and a wider involvement from non-dedicated Observatory-like systems over the coming years.

NEXT STEPS

- *Additional WO installations will be planned for members within the WSTNet eco-systems as well as linking to other academic and community data sources as part of an expansion of the WO template and its resources*
- *The topic of data provenance will need to be further explored and "baked into" WO solutions such that the WO can provide a trusted data from known sources*
- *The ethics of WO's and WO data are already hot research topics and will continue to take a central role in the delivery of systems which balance social insights against social values/norms such as privacy*
- *Legal and Licensing models for original and derived data continue to provide a challenge for big/broad data systems and the topic of data sovereignty will figure particularly prominently for internationally (globally) distributed data systems.*

REFERENCES

The following comprises a selection of literature dealing with the WO, Observatory-related systems and the socio-technical effects and social machines generated by the Web that are the focus of study of the W³O.

BIBLIOGRAPHY

Accomazzi, A. & Dave, R., 2011. Semantic Interlinking of Resources in the Virtual Observatory Era. *arXiv.org*, astro-ph.IM.

Anderson, C., 2010. *The Long Tail. 2004*, Wired Magazine.

Ausubel, J.H., 2011. THE NEW INTERNATIONAL DEEP CARBON OBSERVATORY. *11th Gas Workshop Abstracts*.

Barabasi, A.-L., 1999. Emergence of Scaling in Random Networks. *Science (New York, N.Y.)*, 286(5439), pp.509–512.

Benkler, Y., 2006. *Wealth of networks*,

Berners-Lee, T., 1989. *Information Management: A Proposal* <http://www.w3.org>, History.

Berners-Lee, T. & Fischetti, M., 1999. *Weaving the Web*, Texere Publishing.

Berners-Lee, T., Hall, W. & Hendler, J., 2006a. *A Framework for Web Science* Now Pub, Now Pub.

Berners-Lee, T., Hall, W., Hendler, J., et al., 2006b. Computer science. Creating a science of the Web. *Science (New York, N.Y.)*, 313(5788), pp.769–771.

Bower, J.L. & Christensen, C.M., 1995. Disruptive technologies: catching the wave. *cbred.uwf.edu*.

Brown, I.C., Hall, W. & Harris, L., 2013. From Search to Observation. Available at: <http://www2013.org/companion/p1317.pdf>.

Budavári, T. et al., 2003. SkyQuery -- A Prototype Distributed Query Web Service for the Virtual Observatory. *Astronomical Data Analysis Software and Systems (ADASS) XIII*, 295, p.31.

- Burnap, P. et al., 2014. COSMOS: Towards an integrated and scalable service for analysing social media on demand. *International Journal*
- Carr, N., 2011. *The shallows: What the Internet is doing to our brains*, New York: W.W. Norton.
- Chua, T.S. et al., 2012. NExT: NUS-Tsinghua Center for Extreme Search of User-Generated Content. *MultiMedia, IEEE*, 19(3), pp.81–87.
- Contractor, N., 2009. The emergence of multidimensional networks. *Journal of Computer-Mediated Communication*.
- Difranzo, D. et al., 2014. The web observatory extension: facilitating web science collaboration through semantic markup. In International World Wide Web Conferences Steering Committee.
- Eckert, P., 2006. *Communities of Practice*. (EK Brown, RE Asher, & J. MY Simpson, Eds.) *Encyclopedia of language & linguistics*,
- Ellul, J., 1989. The technological society (1954). *Perspectives on the computer revolution*.
- Engelbart, D.C. et al., 1962. *Augmenting Human Intellect*,
- Gloria, M. & McGuinness, D.L., 2014. Building Web Observatories for Health Web Science. In Proceedings of Web Science
- Hall, W. & De Roure, D., 2009. The evolution of the Web and implications for eResearch. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 367(1890), pp.991–1001.
- Hall, W. & Tiropanis, T., 2012. Web evolution and web science. Available at: <http://www.sciencedirect.com/science/article/pii/S1389128612003581>.
- Hall, W. et al., 2013. The Southampton University Web Observatory. *Multiple values selected*, pp.1–4. Available at: <https://eprints.soton.ac.uk>
- Hall, W. et al., 2014. The Web Science Observatory-The Challenges of Analytics over Distributed Linked Data Infrastructures. *ERCIM*
- Hendler, J., 2013. Broad Data: Exploring the Emerging Web of Data. *Big Data*, 1(1), pp.18–20.
- Hendler, J. & Berners-Lee, T., 2010. From the Semantic Web to social machines: A research

- challenge for AI on the World Wide Web. *Artificial Intelligence*, 174(2), pp.156–161.
- Hendler, J. & Golbeck, J., 2007. *Metcalf's Law, Web 2.0, and the Semantic Web*,
- Hendler, J. et al., 2008. Web science: an interdisciplinary approach to understanding the web. *Commun. ACM* (), 51(7), pp.60–69.
- Holsapple, C.W., Whinston, A.B. & Benamati, J.H., 1996. *Holsapple: Decision support systems: A knowledge-based...* - Google Scholar,
- Kauffman, S.A., 1993. *The Origins of Order: Self-Organization and Selection in Evolution*, Oxford University Press.
- Latour, B., 2005. Reassembling the Social - An Introduction to Actor-Network-Theory. *dss-edit.com*, -1.
- Latour, B., 2008. *Reassembling the Social-an Introduction to Actor-Network-Theory 16th IFOAM Organic World Congress*, Modena.
- Lee, T.B., Hendler, J. & Lassila, O., 2001. The semantic web. *Scientific American*.
- Levine, R., 2009. *The cluetrain manifesto*, New York: Basic Books.
- Malone, T., Laubacher, R. & Dellarocas, C., 2009. Harnessing Crowds: Mapping the Genome of Collective Intelligence. pp.1–20.
- Malone, T.W., Crowston, K. & Herman, G.A., 2003a. Organizing Business Knowledge: The MIT Process Handbook. *Organizing Business Knowledge: The MIT Process Handbook*.
- Malone, T.W., Crowston, K. & Herman, G.A., 2003b. *The MIT process handbook*,
- Mann, R. et al., 2003. Xml in the Virtual Observatory. *Large Telescopes and Virtual Observatory: Visions for the Future*, 8, p.37.
- Martin, U. & Pease, A., 2013. Mathematical practice, crowdsourcing, and social machines. In CICM'13: Proceedings of the 2013 international conference on Intelligent Computer Mathematics. Springer-Verlag.
- Maslow, A. & Herzberg, A., 1954. Hierarchy of needs. *AH Maslow ea*.
- McKelvey, K. & Menczer, F., 2013. Interoperability of Social Media Observatories. pp.1–3.
- McNiff, J., 2013. *Action Research*, Routledge.

- McNiff, J. & Whitehead, J., 2011. *All You Need to Know About Action Research*, SAGE.
- Meira, S.R.L. et al., 2011. The Emerging Web of Social Machines. *IEEE International Computer Software and Applications Conference. Proceedings*, 24(17), pp.26–27.
- Monge, P.R. & Contractor, N.S., 2001. Emergence of communication networks. *The new handbook of organizational*
- O'reilly, T., 2007. What is Web 2.0: Design patterns and business models for the next generation of software. *Communications and Strategies*.
- Page, K.R. & De Roure, D., 2013. Trajectories through Social Machines. pp.1–4.
- Pearson, J. & Shim, J.P., 1994. An empirical investigation into decision support systems capabilities: A proposed taxonomy. *Information & Management*, 27(1), pp.45–57.
- Pinch, T.J. & Bijker, W.E., 1987. The Social Construction of Facts and Artifacts: Or How the Sociology of. ... *in the sociology and history of technology*.
- Power, D.J., 2001. Supporting Decision-Makers: An Expanded Framework. pp.1–6.
- Power, D.J. & Sharda, R., 2007. Model-driven decision support systems: Concepts and research directions. *Decision Support Systems*, 43(3), pp.1044–1061.
- Quinn, P. & Hanisch, B., 2004. The International Virtual Observatory Alliance. *Optimizing scientific return for astronomy through information technologies: 24-25 June, 2004, Glasgow, Scotland, United Kingdom*, 5493, p.137.
- Rousch, W., 2005. Social Machines | MIT Technology Review. Available at: <http://www.technologyreview.com/featuredstory/404466/social-machines/>.
- Schumpeter, J.A., 1935. The Analysis of Economic Change. *The Review of Economics and Statistics*, 17(4), p.2.
- Shadbolt, N.R. et al., 2013. *Towards a classification framework for social machines*, International World Wide Web Conferences Steering Committee.
- Shneiderman, B., 2007. Web science: a provocative invitation to computer science. 50(6), pp.25–27. Available at: http://dl.acm.org/ft_gateway.cfm?id=1247022&type=html.
- Sismondo, S., 2011. *An Introduction to Science and Technology Studies - Sergio Sismondo - Google Books*
- Smart, P.R., Simperl, E. & Shadbolt, N., 2014. A Taxonomic Framework for Social Machines.

Available at: <http://eprints.soton.ac.uk/362359/>.

Smith, M.R. & Marx, L., 1994. *Does Technology Drive History*, MIT Press (MA).

Tarte, S.M., De Roure, D. & Willcox, P., 2014. *Working out the plot: the role of stories in social machines*, International World Wide Web Conferences Steering Committee. Available at: <http://dl.acm.org/citation.cfm?id=2567948.2578839>.

Tinati, R. & Carr, L., 2012. Understanding Social Machines. In *Privacy, Security, Risk and Trust (PASSAT)*, 2012 International Conference on and 2012 International Conference on Social Computing (SocialCom. IEEE, pp. 975–976.

Tinati, R. et al., 2013. The HTP model: understanding the development of social machines. ... *the 22nd international conference*

Tiropanis, T. & Hall, W., 2013. The Web Science Observatory. Available at: <http://eprints.soton.ac.uk/354604/1/TheWebScienceObservatory-postprint.pdf>.

Tiropanis, T., Rowland-Campbell, A. & Hall, W., 2014a. *Government as a social machine in an ecosystem*, International World Wide Web Conferences Steering Committee.

Tiropanis, T., Wang, X., et al., 2014b. Building a Connected Web Observatory: Architecture and Challenges.

Van Kleek, M. et al., 2014. 7 billion home telescopes: observing social machines through personal data stores. In *WWW Companion '14: Proceedings of the companion publication of the 23rd international conference on World wide web companion*.

Van Kleek, M., Smith, D.A. & Hall, W., 2013. "The Crowd Keeps Me in Shape": Social Psychology and the Present and Future of Health Social Machines. pp.1–5.

Wang, H. et al., 2013. Semantically-enabled Knowledge Discovery in the Deep Carbon Observatory. *AGU Fall Meeting*

Wiener, N., 1949. *Cybernetics, Or Control and Communication in the Animal and the Machine*. Norbert Wiener,

Woolley, A.W.A. et al., 2010. Evidence for a collective intelligence factor in the performance of human groups. *Science (New York, N.Y.)*, 330(6004), pp.686–688.

Yip, M. & Webber, C., 2012. Structural analysis of online criminal social networks. In *2012 IEEE International Conference on Intelligence and Security Informatics (ISI 2012)*. IEEE, pp. 60–65.

List of symbols, abbreviations and acronyms

BWOW	Building Web Observatories Workshop
SUWO	Southampton Uni Web Observatory
Web Observatory	Virtual instrument for studying data on/about the Web
W³O	World-wide web observatory (the)
WO	Web Observatory (a)
WOW	Web Observatory Workshop
WST	Web Science Trust
WSTNet	WST network of labs
