ERC Consolidator Grant 2014

Annex 1 to the Grant Agreement
(Description of the Action)
Part B

Proposal Acronym:	**COMPROP**
Proposal number:	**648311**
Proposal Title:	**Computational Propaganda: Investigating the Impact of Algorithms and Bots on Political Discourse in Europe**
Principal Investigator:	**Philip N. Howard**
Host Institution:	**Oxford University**

Table of Contents

1. a) Summary 3
1. b) Curriculum vitae – Funding ID 4
1. c) Summary 7
2. a) State of the Art and Objectives 13
2. b) Methodology 19
2. c) Resources 28

1. a) Summary

Social media can have an impressive impact on civic engagement and political discourse. Yet increasingly we find political actors using digital media and automated scripts for social control. Computational propaganda—through bots, botnets, and algorithms—has become one of the most concerning impacts of technology innovation. Unfortunately, bot identification and impact analysis are among the most difficult research challenges facing the social and computer sciences.

COMPROP objectives are to advance a) rigorous social and computer science on bot use, b) critical theory on digital manipulation and political outcomes, c) our understanding of how social media propaganda impacts social movement organization and vitality. This project will innovate through i) "real-time" social and information science actively disseminated to journalists, researchers, policy experts and the interested public, ii) the first detailed data set of political bot activity, iii) deepened expertise through cultivation of a regional expert network able to detect bots and their impact in Europe.

COMPROP will achieve this through multi-method and reflexive work packages: 1) international qualitative fieldwork with teams of bot makers and computer scientists working to detect bots; 2a) construction of an original event data set of incidents of political bot use and 2b) treatment of the data set with fuzzy set and traditional statistics; 3) computational theory for detecting political bots and 4) a sustained dissemination strategy. This project will employ state-of-the-art "network ethnography" techniques, use the latest fuzzy set / qualitative comparative statistics, and advance computational theory on bot detection via cutting-edge algorithmic work enhanced by new crowd-sourcing techniques.

Political bots are already being deployed over social networks in Europe. COMPROP will put the best methods in social and computer science to work on the size of the problem and the possible solutions.

Behind the bots are norms and innovation networks. By mapping out such networks we can get closer to understanding modes of operation and design trajectories for bot builders. The PI has a demonstrated record of scholarship involving socio-technical systems of hackers, computer scientists, and innovation networks. Moreover, COMPROP work packages actively engage with teams of computer scientists involved in bot detection, the fieldwork will be with the authors of automated scripts, and the budget specifically provides post-doc opportunities for computer scientists interested in the political impact of bots. Accordingly, the primary domain is SH2 Social Sciences & Humanities - The Social World, Diversity and Common Ground, but the valuable second domain is PE6 Physical Sciences & Engineering - Computer Science and Informatics.

1. b) Curriculum Vitae – Funding ID

PERSONAL INFORMATION

Family name, First name:	Howard, Philip
Researcher unique identifier(s)	EX2014D199212
Date of birth:	December 9, 1970
URL for web site:	http://www.philhoward.org

- **EDUCATION**

2002 PhD, Sociology Department, Northwestern University, Evanston, United States
1996 Certificate, United Nations Civilian Training for International Peacekeeping Missions, Scuola Superiore Sant'Anna, Pisa, Italy
1994 MSc, Economics, London School of Economics, London, United Kingdom
1993 BA, Political Science, University of Toronto, Toronto, Canada

- **CURRENT POSITION(S)**

2015– Professor, Oxford Internet Institute, Oxford University, United Kingdom
2012– Non-Resident Fellow, Tow Digital Journalism Center, Columbia University, United States

- **PREVIOUS POSITIONS**

2013–2015 Professor and Director, Center for Data, Media and Society, School of Public Policy, Central European University, Hungary
2012–2013 Professor, Center for Information Technology Policy, Woodrow Wilson School of International and Public Affairs, Princeton University, United States
2009 Visiting Professor, Comparative Social Science Studies, University of Oslo, Norway
2002–2012 Assistant, Associate and Full Professor, Department of Communication, Information School and Jackson School of International Studies, University of Washington, United States

- **FUNDING ID SUMMARY**

Prior support from the National Science Foundation (NSF) in the US (IIS-0713074, €249,000, 2007-2010 and ITR-0326101, €872,000 million, 2003–2005) allowed the PI to investigate new ways of measuring the impact of the engineering standards setting process on technology diffusion in Central Asia, Tanzania, and the Middle East. PI Howard's "RAPID - Social Computing and Political Transition in Tunisia," (IIS-1144286, €33,300, 2011) allowed for additional fieldwork in Tunisia during that country's first real election and their simultaneous efforts at developing open telecommunications standards and less restrictive information policy. Coupled with support from the Knight Foundation and fully funded fellowship appointments at Stanford and Princeton, these early achievements have enabled PI Howard to become one of the foremost experts on the political and policy process that can turn innovations in computer science and engineering into tools for civic engagement or social control. NSF support has been acknowledged at in multiple articles and research monographs, including *Digital Origins of Dictatorship and Democracy: Information Technology and Political Islam* (New York: Oxford University Press 2010) which is acknowledged to be the most prescient work on the role of ICTs in the Arab Spring and *Democracy's Fourth Wave? Digital Media and the Arab Spring* (New York: Oxford University Press, 2013) the first major manuscript on the role of social media in those popular uprisings.

Over the last 10 years the PI has demonstrated an ability to propose and conduct ground-breaking research, with a record of awards from diverse funding bodies:

- ✶ public agencies such as the Canadian Federal Privacy Commissioner and US Institute of Peace;
- ✶ public scientific agencies such as the US NSF, Canada's Social Science and Humanities Research Council (SSHRC), and the UK's Economic and Social Research Council (ESRC);
- ✶ private research funders like the Knight Foundation;
- ✶ industry research labs like Microsoft Research and Intel's People and Practices Research Group.

Aside from the salary and benefits awards from full residential fellowships to Princeton and Stanford, to date the PI has managed €1.5 million from funders. His first large grant management experience was as co-PI of a

€872,000 grant from the NSF to study technology diffusion in five Central Asian countries. Since then he has applied for and been awarded another €678,000 in support for projects on information and intelligent systems, technology diffusion and democratization, and political communication. Approximately €124,000 of that amount has been awards the PI sought to support students doing graduate projects in conjunction with my broad research agenda—the PI actively supports and mentors the next generation of researchers.

- **SUPERVISION OF GRADUATE STUDENTS AND POSTDOCTORAL FELLOWS**

The PI has a strong record of mentoring students through degree completion and publishing:

- ✶ Master Students = 14. Many of these students successfully defended their projects and went on to PhD programs. Additionally I have mentored 12 of them by publishing with them as co-authors.
- ✶ PhD Students = 15. Of these, 10 have successfully defended their projects and accepted tenure-track positions at prominent universities in Europe (Universities of Amsterdam, Bournemouth, Glasgow), North America (Universities of Michigan, North Carolina-Chapel Hill, Illinois-Chicago, New Mexico, and American University), around the world (Gulf University for Science and Technology, National University of Singapore), and in Industry research centres (Microsoft Research, intel Research).
- ✶ Post Docs = 1. I supervised a post-doctoral fellow who recently accepted a tenure track appointment at Concordia University in Canada.

- **TEACHING ACTIVITIES**

2014 Founding Professor, School of Public Policy, Central European University, Hungary
2012–13 Visiting Faculty–Digital Media and Politics, Columbia University, United States
2009 Visiting Faculty–Comparative Information Societies, Oslo University, Norway
2002–2012 Assistant, Associate, and Full Professor–Communication and Media Studies, University of Washington, United States

- **INSTITUTIONAL RESPONSIBILITIES**

2013– Founding Professor, School of Public Policy, Central European University, Hungary
2013– Director, Center for Data, Media and Society, School of Public Policy, Central European University, Hungary
2002–2012 Graduate Student Advisor, University of Washington, United States
2007–2008 Member of the Faculty Senate, University of Washington, United States

- **COMMISSIONS OF TRUST**

2011–2013 Best Dissertation Award Committee, American Sociological Association
2010– Editorial Board and Reviewer, Journal of Communication, Information, Communication & Society, The Information Society, Journal of Latin American Urban Studies, Journal of Information Technology and Politics
2010– Review Panel Member, NSF and Canadian SSHRC
2008– Reviewer, Oxford, Cambridge, and MIT University Presses
2002– Reviewer, American Journal of Sociology, Comparative Politics, Journal of Communication, Journal of Computer-Mediated Communication, The Information Society, Information, Communication and Society, New Media & Society, Political Communication, Public Opinion Quarterly

- **MEMBERSHIPS OF SCIENTIFIC SOCIETIES**

2012– Member, "Microsoft Social Media Collective," Microsoft Research, Boston, United States
2002– Association of Internet Researchers, American Political Science Association, American Sociological Association, International Communication Association, International Studies Association

- **MAJOR COLLABORATIONS**

- ✶ With Dr. Muzammil Hussain, Communication Studies, University of Michigan I have written several articles for leading peer-review journals in international affairs and internet studies on technology diffusion and political norms. He was my first PhD student, and we have co-edited *State Power 2.0: Digital Networks and Authoritarian Rule* (London, UK: Ashgate, 2013) and co-authored *Democracy's Fourth Wave? Digital Media and the Arab Spring* (New York: Oxford University Press, 2013).

- ✶ With Dr. Andrew Chadwick, Department of Politics and International Relations, UCL-Royal Holloway I have edited the field-defining *Handbook of Internet Politics*. London: Routledge, 2009.
- ✶ With Dr. Monroe Price, Center for Global Communication, Annenberg School, University of Pennsylvania I have multiple collaborations through our research centres, including student exchanges, an intensive PhD summer school on information technology and civic engagement, and faculty exchanges.

Appendix: All on-going and submitted grants and funding of the PI (Funding ID)

On-going Grants

Project Title	*Funding source*	*Amount (Euros)*	*Period*	*Role of the PI*	*Relation to current ERC proposal*
The Production/ Detection of Bots	NSF, United States	160,000	September 2014–August 2016	PI	Smaller US Component supporting Seattle-area interviews.

Applications

Project Title	*Funding source*	*Amount (Euros)*	*Period*	*Role of the PI*	*Relation to current ERC proposal*
None					

There is and there will be no funding overlap with the ERC grant requested and any other source of funding for the same activities and costs that are foreseen in this project.

1. c) Summary

- **REPRESENTATIVE PUBLICATIONS (GOOGLE SCHOLAR CITATION COUNTS)**

1. Monographs
   a. Howard, Philip N. and Muzammil Hussain. *Democracy's Fourth Wave? Digital Media and the Arab Spring*. New York: Oxford University Press, 2013. (Citations = 26, published scholarly reviews = 3)
   b. Howard, Philip N. *The Digital Origins of Dictatorship and Democracy: Information Technology and Political Islam*. New York: Oxford University Press, 2011. (Citations = 206, published scholarly reviews = 7, scientific book awards = 1)
   c. Chadwick, Andrew, and Philip N. Howard, eds. *Handbook of Internet Politics*. London: Routledge, 2009. (Citations = 563, Published scholarly reviews = 3)
   d. Howard, Philip N. *New Media Campaigns and the Managed Citizen*. New York: Cambridge University Press, 2006. (Citations = 234, published scholarly reviews = 10, scientific book awards = 2).
   e. Howard, Philip N., and Steve Jones, eds. *Society Online: The Internet in Context*. Thousand Oaks, CA: Sage, 2004. Also published in Spanish as Howard, Philip N., and Steve Jones, eds. *Sociedad on-Line*. Barcelona: Editorial UOC, 2005. (Citations = 715, published scholarly reviews = 6)
2. Major International Peer-Reviewed Multi-Disciplinary Scientific Journals
   a. Howard, Philip N. Special Editor "Social Media and Political Change: Capacity, Constraint, and Consequence." *Journal of Communication 62*, no. 2 (2012). (Citations = 379)
   b. Howard, Philip N., Lee Rainie and Steve Jones, "Days and Nights on the Internet: The Impact of a Diffusing Technology," *American Behavioral Scientist* 45, November 2001, pp. 382-404. (Citations = 563)
3. Leading International Peer-Reviewed Journals
   a. Howard, Philip N. "Participation, Civics and Your Next Coffee Maker." *Policy & Internet* 6, no. 2 (2014): forthcoming. (Citations = N/A)
   b. Howard, Philip N. and Muzammil Hussain. "The Role of Digital Media." *Journal of Democracy* 22, no. 3 (2011): 35-48. (Citations = 113)
   c. Howard, Philip N. "Network Ethnography and the Hypermedia Organization: New Media, New Organizations, New Methods." *New Media & Society* 4, no. 4 (2002): 550–74. (Citations = 217)

- **INVITED PRESENTATIONS**

1. Recent Keynote Lectures at Scientific Meetings
   a. "Digital Activism in the Developing World," MacArthur Research Network on Youth and Participatory Politics, Istanbul, February 2013.
   b. "Social Media and Digital Diplomacy," Canadian Ministry of Foreign Affairs, March 2013.
   c. "Scenarios for Internet Governance," International Governance Innovation, CIGI, April 2013.
   d. "Digital Media and the Arab Spring," Harvard University, Radcliffe Institute, March 2012.
2. Presentations at Peer Reviewed, International Established Conferences
   a. Philip N. Howard, "Digital Media, Democracy, and Dictators," International Studies Association, April 2011.
   b. Philip N. Howard, Panel Chair, "Digital Media Power Struggles: Contentious Politics and Social Media Mobilization," International Studies Association, April 2011.
   c. Philip N. Howard, Muzammil Hussain, and Sheetal Agarwal. "When Do States Disconnect Their Digital Networks?" American Political Science Association, Seattle, September 2011.
   d. Philip N. Howard, "Digital Media and Discontent: A Fuzzy Look at the Arab Spring," International Communication Association, Boston MA, May 2011.
   e. Philip N. Howard, Panel Chair, "Religion, Technology, and Transformations in State and Society Relations," American Political Science Association, Washington DC, September 2010.
   f. Philip N. Howard, "The Internet and Islam - The Digital Origins of Dictatorship and Democracy," American Political Science Association, Washington DC, September 2010.
   g. Philip N. Howard and Muzammil Hussain, "Information Technology and Democratic Islam," APSA Political Communication Pre-Conference, Washington DC, September 2010.
   h. Philip N. Howard, "Is the Internet Redefining the Concept of Citizenship?", American Sociological Association, Atlanta, August 2010.

i. Philip N. Howard, Panel Chair, "The Challenge of Literacy in an Information Society", International Communication Association, Singapore, July 2010.
j. Philip N. Howard, "Information Technology and Democratic Islam", International Communication Association, Singapore, July 2010.

3. Invited Lectures at International Advanced Schools
   a. "The Pax Technica" - Oxford University, April 2014; Central European University, Budapest, February 2014; Stanford University, February 2013.
   b. "The New Cold Media War" - Princeton University, CITP, October 2013; University of Pennsylvania, Annenberg School, January 2013.
   c. "Democracy's Fourth Wave? Digital Media and the Arab Spring" - National Democracy Institute, April, 2013; US Institute of Peace, April 2013; George Washington University, April 2013; Moscow State Humanities University, July 2012; Freie Universitat Berlin, June 2012; Princeton University, CITP, April 2012.

- **PRIZES AND AWARDS**

2013 Fellow, Microsoft Research, Cambridge, MA, United States. Full resident fellowship award with stipend and research support, 1 month.

2012–2013 Fellow, Center for Information Policy, Princeton University, United States. Full resident fellowship award with stipend and research support, 18 months.

2011 Best Book Award, Information Technology and Politics Section, American Political Science Association. Competitively awarded honor for best scientific book published that year by a member of that professional section.

2008–2009 Fellow, Center for Advanced Study in Behavioral Sciences, Stanford University, United States. Full resident fellowship award with stipend and research support, 12 months.

2008 Outstanding Book Award, International Communication Association. Competitively awarded honor for best scientific book published that year by the entire professional association of the discipline.

2006 Best Book Award, Communication Technology & Society Section, American Sociological Association. Competitively awarded honor for best scientific book published that year by a member of that professional section.

2003–2004 Fellow, Stanhope Centre for Communications Policy Research, London School of Economics, United Kingdom. Full resident fellowship award with stipend and research support, 6 months.

- **OTHER KEY INDICATORS OF ACHIEVEMENT**

★ Recently guest edited a special issue of the major disciplinary flagship journal in media and communication studies, the *Journal of Communication*, on social media and political change (Volume 64 Issue 2). Along with cultivating and editing manuscripts, wrote the framing essay "Social Media and Political Change: Capacity, Constraint, and Consequence." (Citations = 379)

★ He is one of the few social science scholars to win book awards from professional associations across three disciplines: political science (APSA), sociology (ASA), and communication (ICA).

★ He has had multiple high profile residential fellowships from CASBS at Stanford, CITP at Princeton, and the Tow Center at Columbia.

★ He has managed over €1.5 million in externally funded research projects from public agencies (NSF, USIP), private foundations (Knight, Open Society Foundation), and industry (Microsoft, Intel).

★ He has published 8 books, 24 peer-review articles (many in major and leading international peer-reviewed journals), 18 book chapters, 52 conference papers, and 12 policy papers with funders and think tanks.

★ His research monographs have been critically acclaimed in 31 review essays in scholarly journals.

★ His research has been prominently featured in news media around the world, including *The Atlantic*, *Slate*, *The Guardian* and *Der Spiegel*.

★ The PI has formally published 3 datasets that have been downloaded hundreds of times and been used in dozens of research articles.

- Howard, Philip, Laura Busch, and Spencer Cohen. "ICT Diffusion and Distribution Dataset, 1990–2007." ICPSR Public Dataset #9908 (2008, with 2010 revision).
- Howard, Philip, Mary Joyce, and Frank Edwards, "Global Digital Activism Dataset" ICPSR Public Dataset #33871 (2012).

- Howard, Philip N., Muzammil Hussain, and Sheetal Agarwal. States Interference with Digital Networks, 2000-2012. World Information Access Project. 2012.

References Used in Extended Synopsis and Full Scientific Proposal

Abbott, Andrew Delano. 1988. *The System of Professions : An Essay on the Division of Expert Labor*. Chicago: University of Chicago Press.

Aday, Sean, Henry Farrell, Marc Lynch, John Sides, and Deen Freelon. 2012. "Blogs and Bullets II: New Media and Conflict after the Arab Spring". United States Institute of Peace. http://www.usip.org/publications/blogs-and-bullets-ii-new-media-and-conflict-after-the-arab-spring.

Aday, Sean, Henry Farrell, Marc Lynch, John Sides, John Kelly, and Ethan Zuckerman. 2010. "Blogs and Bullets: New Media in Contentious Politics". Washington, DC: USIP. http://www.usip.org/publications/blogs-and-bullets-new-media-in-contentious-politics.

Antin, Judd, and Aaron Shaw. 2012. "Social Desirability Bias and Self-Reports of Motivation: A Study of Amazon Mechanical Turk in the US and India." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2925–34. CHI '12. New York, NY, USA: ACM. doi:10.1145/2207676.2208699. http://doi.acm.org/10.1145/2207676.2208699.

Bennett, W. Lance, and Alexandra Segerberg. 2013. *The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics*. Cambridge University Press. http://books.google.com/books?hl=en&lr=&id=nZFtAAAAQBAJ&oi=fnd&pg=PR8&dq=lance+bennett+connective+action&ots=RoGfU91OUi&sig=2CT0w7r-xC_VbTAcrZY1VdUFqLM.

Beyer, Jessica Lucia. 2014. *Expect Us: Online Communities and Political Mobilization*. New York, NY: Oxford University Press.

Bijker, Wiebe E., Thomas Parke Hughes, and T. J. Pinch. 1987. *The Social Construction of Technological Systems : New Directions in the Sociology and History of Technology*. Cambridge, Mass.: MIT Press.

Burrell, Jenna. 2012. *Invisible Users: Youth in the Internet Cafés of Urban Ghana*. Cambridge, Mass.: The MIT Press.

Carrieri, Matthew, Ronald Deibert, and Saad Omar Khan. 2014. "Information Infrastructure and Anti-Regime Protests in Iran and Tunisia." In *State Power 2.0: Authoritarian Entrenchment and Political Engagement Worldwide*, edited by Philip N. Howard and Muzammil M. Hussain, 45–56. Farnham, UK: Ashgate.

Castells, Manuel. 2009. *Communication Power*. New York, NY: Oxford University Press.

Chu, Zi, Steven Gianvecchio, Haining Wang, and Sushil Jajodia. 2010. "Who Is Tweeting on Twitter: Human, Bot, or Cyborg?" In *Proceedings of the 26th Annual Computer Security Applications Conference*, 21–30. ACM. http://dl.acm.org/citation.cfm?id=1920265.

Coleman, E. Gabriella. 2013. *Coding Freedom: The Ethics and Aesthetics of Hacking*. Princeton University Press. http://books.google.com/books?hl=en&lr=&id=O42iNERgx7oC&oi=fnd&pg=PP2&dq=coding+freedom&ots=BNUHtEbMlm&sig=v6vNZtpdPL9GcIzdPNw-VijhFCQ.

Deibert, Ronald. 2012. *Access Contested Security, Identity, and Resistance in Asian Cyberspace Information Revolution and Global Politics*. Cambridge, MA: MIT Press. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=421819.

Diamond, Larry. 2010. "Liberation Technology." *Journal of Democracy* 21 (3): 69–83.

Downey, John, and James Stanyer. 2010. "Comparative Media Analysis: Why Some Fuzzy Thinking Might Help. Applying Fuzzy Set Qualitative Comparative Analysis to the Personalization of Mediated Political Communication." *European Journal of Communication* 25 (4): 331–47. doi:10.1177/0267323110384256.

Earl, Jennifer, and Katrina Kimport. 2011. *Digitally Enabled Social Change: Activism in the Internet Age*. 1st ed. The MIT Press.

Edwards, F., Philip N. Howard, and Mary Joyce. 2013. "Digital Activism and Non-Violent Conflict". Digital Activism Research Project.

Elliott, Chris. 2014. "The Readers' Editor On… pro-Russia Trolling below the Line on Ukraine Stories." *The Guardian*, May 4, sec. Comment is free. http://www.theguardian.com/commentisfree/2014/may/04/pro-russia-trolls-ukraine-guardian-online.

Haas, Peter M. 1992. "Introduction: Epistemic Communities and International Policy Coordination." *International Organization* 46 (1): 147–286.

Howard, Philip N. 2002. “Network Ethnography and the Hypermedia Organization: New Media, New Organizations, New Methods.” *New Media & Society* 4 (4): 550–74.

———. 2006. “Introduction: The Hypermedia Campaign.” In *New Media Campaigns and the Managed Citizen*, xxii, 264 p. New York: Cambridge University Press. http://www.loc.gov/catdir/toc/ecip059/2005008088.html.

———. 2010. *The Digital Origins of Dictatorship and Democracy: Information Technology and Political Islam*. Oxford University Press. http://books.google.com/books?hl=en&lr=&id=T6xoAgAAQBAJ&oi=fnd&pg=PR9&dq=digital+origins+of+dictatorship&ots=_fUkT7Mqzz&sig=luGx7448iwAz1c1dYY1vC8sgWtg.

Howard, Philip N., and Muzammil M. Hussain. 2013. *State Power 2.0: Authoritarian Entrenchment and Political Engagement Worldwide*.

insidecroydon. 2014. “Jasper Admits to Using Twitter Bots to Drive Election Bid.” *Inside Croydon*. Accessed May 10. http://insidecroydon.com/2012/11/26/jasper-admits-to-using-twitter-bots-to-drive-election-bid/.

Karlova, Natascha A., John L. Marino, Peyina Lin, and Michael B. Eisenberg. 2014. “Challenges and Opportunities in Virtual Worlds: Informed Consent and Data Collection.” March 1. http://hdl.handle.net/2142/47282.

Kim, Won, Ok-Ran Jeong, Chulyun Kim, and Jungmin So. 2010. “On Botnets.” In *Proceedings of the 12th International Conference on Information Integration and Web-Based Applications & Services*, 5–10. ACM. http://dl.acm.org/citation.cfm?id=1967488.

Kittur, Aniket, Jeffrey V. Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. 2013. “The Future of Crowd Work.” In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, 1301–18. CSCW ’13. New York, NY, USA: ACM. doi:10.1145/2441776.2441923. http://doi.acm.org/10.1145/2441776.2441923.

Krinsky, John, and Ann Mische. 2013. “Formations and Formalisms: Charles Tilly and the Paradox of the Actor.” *Annual Review of Sociology* 39 (1): 1–26. doi:10.1146/annurev-soc-071312-145547.

Latour, Bruno. 2005. *Reassembling the Social: An Introduction to Actor Network Theory*. Oxford: Oxford University Press.

MacKinnon, Rebecca. 2012. *Consent of the Networked: The Worldwide Struggle For Internet Freedom*. Basic Books.

Maher, Katherine. 2011. “What’s Happening in Tunisia?” *NDITech:* http://www.demworks.org/blog/2011/01/whats-happening-tunisia.

Morozov, Evgeny. 2011. *The Net Delusion: The Dark Side of Internet Freedom*. PublicAffairs.

Neff, Gina. 2011. *Venture Labor*. Cambridge, MA: MIT Press.

Podolny, Joel, and Karen Page. 1998. “Network Forms of Organization.” *Annual Review of Sociology* 24: 57–76.

Qtiesh, Anas. 2011. “Spam Bots Flooding Twitter to Drown Info About #Syria Protests.” *Global Voices Advocacy*. http://advocacy.globalvoicesonline.org/2011/04/18/spam-bots-flooding-twitter-to-drown-info-about-syria-protests/.

Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. 1st ed. University Of Chicago Press.

Rihoux, Benoît, and Charles C. Ragin. 2008. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Sage Publications, Inc.

Rogers, Richard. 2013. *Digital Methods*. Cambridge, Massachusetts: The MIT Press.

Rosenberg, Yuval. 2013. “62 Percent of All Web Traffic Comes from Bots.” *The Week*. December 16. http://theweek.com/article/index/254183/62-percent-of-all-web-traffic-comes-from-bots.

Saxenian, AnnaLee. 1996. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.

Shirky, Clay. 2008. *Here Comes Everybody: The Power of Organizing without Organizations*. Penguin. http://books.google.com/books?hl=en&lr=&id=UNxU-2s2sQYC&oi=fnd&pg=PT12&dq=shirky&ots=60XETuWWMB&sig=mPxJQsOgJgD5jb0pmibXWzwADig.

Van der Windt, Peter. 2014. “From Crowdsourcing to Crowdseeding: The Cutting Edge of Empowerment?” In *Bits and Atoms*, edited by Steven Livingston and Gregor Walter-Drop, 144–55. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199941599.001.0001/acprof-9780199941599-chapter-10.

Wagstaff, Keith. 2013. “1 in 10 Twitter Accounts Is Fake, Say Researchers - NBC News.com.” *NBC News*. November 26. http://www.nbcnews.com/technology/1-10-twitter-accounts-fake-say-researchers-2D11655362.

Woolley, Samuel, and Philip N. Howard. forthcoming. “Social Media, Revolution, and the Rise of the Political Bot.” *Routledge Handbook of Media, Conflict, and Security*

Young, Oran. 1991. “Political Leadership & Regime Formation: On the Development of Institutions in International Society.” *International Organization* 43 (4): 349–75.

2. a) State of the Art and Objectives

Introduction

Social media has revolutionized the way people discuss current affairs and obtain political news and information. Contact with social media helps young people cultivate a political identity and engage civically in both authoritarian and democratic regimes.[1] Activist causes and democratic movements have been born, organized and disseminated on sites like Facebook, Twitter, Weibo, and YouTube.[2] Yet another, more propagandistic, side of this new technology exists.[3] Security experts find that over 10 percent of content on social media websites, and 62 percent of web traffic, is bot generated.[4] This project on computational propaganda (COMPROP) will build a team of researchers under the leadership of Principal Investigator (PI) Howard to do: 1) international fieldwork (participant observation, interviews) with teams of bot makers and the people building bot detectors; 2a) construction of an original event-dataset of known incidents of bot use by political actors; 2b) treatment of the original data set with both traditional statistical techniques and state-of-the-art fuzzy logic statistics and 3) computational theory about how bots can be detected during sensitive moments when the manipulation of public opinion would have detrimental effects on public life.

The word "botnet" comes from combining "robot" with "network." It describes a collection of programs that communicate across multiple devices to perform some task. The tasks can be simple and annoying, like generating spam. The tasks can be aggressive and malicious, like choking off exchange points, or launching denial-of-service attacks. And not all are developed to advance political causes. Some seem to have been developed for fun or to support criminal enterprises, but all share the property of deploying messages and replicating themselves.[5] Chu et al. distinguish two types of bots on Twitter: legitimate and malicious. Legitimate bots generate a large amount of benign tweets that deliver news or update feeds. Malicious bots, on the other hand, spread spam by delivering appealing text content with the link directed to spam or malicious content.[6] Botnets are created for many reasons: spam, DDoS attacks, theft of confidential information, click fraud, cyber sabotage, and cyber warfare. According to Kim et al., many governments have been strengthening their cyber warfare capabilities for both defensive and offensive purposes. In addition, political actors and governments worldwide have begun using bots to manipulate public opinion, choke off debate, and muddy political issues.[7]

COMPROP is a multi-method project that will research the use of "bots", botnets and other automated scripts for social control and public opinion manipulation, using an innovative combination of ethnography, social network analysis, and fuzzy set logic. The fieldwork will provide us with the insights to understand the techniques of bot creators. After gaining insight from the inner workings of bot labs, and the processes used by bot detectors, the team will start developing a prototype of a crowd sourced bot detection system that can practically help identify abuse of the public trust, let us practically help improve the flow of traffic in social computing, and let us advance theory about the role of new technologies in contemporary political communication. The team will develop original computational theory about how political bots can be improved, via cutting edge-algorithmic work enhanced by new crowd-sourcing techniques.

COMPROP will help policy makers better understand the relationship between computational propaganda and political processes, and it will help bot detectors understand the evolution of automated scripts for manipulating social networks. Bots have already been employed by political candidates in European

---

[1] W. Lance Bennett and Alexandra Segerberg, *The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics*, 2013; Philip N. Howard, *The Digital Origins of Dictatorship and Democracy: Information Technology and Political Islam* (New York, NY: Oxford University Press, 2010).

[2] F. Edwards, Philip N. Howard, and Mary Joyce, "Digital Activism and Non-Violent Conflict" (Digital Activism Research Project, 2013); Jennifer Earl and Katrina Kimport, *Digitally Enabled Social Change* (The MIT Press, 2011).

[3] Samuel Woolley and Philip N. Howard, "Social Media, Revolution, and the Rise of the Political Bot," *Routledge Handbook of Media, Conflict, and Security*, forthcoming; Larry Diamond, "Liberation Technology," *Journal of Democracy* 21, no. 3 (2010): 69–83; Clay Shirky, *Here Comes Everybody: The Power of Organizing without Organizations* (Penguin, 2008).

[4] Yuval Rosenberg, "62 Percent of All Web Traffic Comes from Bots," *The Week*, December 16, 2013, http://theweek.com/article/index/254183/62-percent-of-all-web-traffic-comes-from-bots.

[5] Won Kim et al., "On Botnets," in *Proceedings of the 12th International Conference on Information Integration and Web-Based Applications & Services* (ACM, 2010), 5–10, http://dl.acm.org/citation.cfm?id=1967488; Keith Wagstaff, "1 in 10 Twitter Accounts Is Fake, Say Researchers - NBC News.com," *NBC News*, November 26, 2013, http://www.nbcnews.com/technology/1-10-twitter-accounts-fake-say-researchers-2D11655362.

[6] Zi Chu et al., "Who Is Tweeting on Twitter: Human, Bot, or Cyborg?," in *Proceedings of the 26th Annual Computer Security Applications Conference* (ACM, 2010), 21–30, http://dl.acm.org/citation.cfm?id=1920265.

[7] Kim et al., "On Botnets."

elections and recently been deployed to shape content exposure on a major UK newspaper.[8] This project will expand the understanding of how this is done, and advance the conversation among researchers in the computing, engineering, and social sciences about the size of the problem and the possible solutions. Moreover, it will generate computational theory for bot detection when the manipulation of public opinion would have detrimental effects on public life in Europe.

Social bots are particularly prevalent on Twitter. They are computer generated programs that post, tweet, or message of their own accord. Often bot profiles lack basic account information such as screen-names or profile pictures. Such accounts have become known as "Twitter eggs" because default profile pictures on the social media site ubiquitously feature an egg.[9] While social media users access from front-end websites, bots get access to such websites directly through a mainline, code-to-code, connection, mainly, through the site's wide-open application programming interface (API), posting and parsing information in real time. Bots are versatile, cheap to produce, and ever evolving. "These bots," argues Dubbin, "whose DNA can be written in almost any modern programming language, live on cloud servers, which never go dark and grow cheaper by day."[10] Over the last two decades, with a rapid increase in just the last two years, developers and their employers have begun to deploy bots beyond mundane commercial tasks like spamming or scraping sites like eBay for bargains. Bots are the primary applications used in carrying out distributed denial of service and virus attacks, email harvesting, and website content theft. Beyond this, and central to the research, lies the explicitly political usage of bots as social media propaganda tools.

The use of political bots varies across regime types. As a preliminary exercise the PI will construct a testable typology of worldwide political bot-usage. The current understandings, based upon initial pilot research, suggest that political bots tend to be used for distinct purposes during three primary events: elections, spin control during political scandals, and national security crises. The usage of bots during these situations extends from the nefarious cause of demobilizing political opposition followers to the *seemingly* innocuous task of padding political candidates' social media "follower" lists. Bots are additionally used to drown-out oppositional or marginal voices, halt protest, and relay "astroturf" messages of false governmental support. Political actors use them in general attempts to manipulate and sway public opinion. It is clear that understanding the creation and usage of this technology is central to generating political equality both on and off line and in fostering genuine advancement of democratic social media possibilities. The extended research of COMPROP will greatly illuminate the process and impact of political bot creation and usage on European infrastructure and political life.

1 b) The Proliferation of Bots and Computational Propaganda

What triggered this COMPROP proposal was an interest in the computational propaganda being carried out by what the PI suspects to be programmers employed by The Syrian Electronic Army (SEA), a hacker network that supports the Syrian government. The group developed a botnet that generates pro-regime content with the aim of flooding the Syrian revolution hashtags (for example #Syria, #Hama, #Daraa) and overwhelming the pro-revolution discussion on Twitter and other social media portals.[11] As the Syrian blogger Anas Qtiesh writes, "These accounts were believed to be manned by Syrian Mokhabarat (intelligence) agents with poor command of both written Arabic and English, and an endless arsenal of bite and insults."[12]

Differing forms of bot generated computational propaganda have been deployed in several other countries: Russia, Mexico, China, Australia, the United Kingdom, the United States, Azerbaijan, Iran, Bahrain, South Korea, and Morocco. Current contemporary political crises in the Thailand, Turkey, and the ongoing situation in Ukraine are seeing the emergence of computational propaganda. Table 1 presents a casual sampling of the diversity of regime types and bot producers around the world, with a democracy score

[8] Chris Elliott, "The Readers' Editor On… pro-Russia Trolling below the Line on Ukraine Stories," *The Guardian*, May 4, 2014, sec. Comment is free, http://www.theguardian.com/commentisfree/2014/may/04/pro-russia-trolls-ukraine-guardian-online.
[9] Chris Taylor, "Newt Gingrich's Twitter Followers Are 8% Human [INFOGRAPHIC]," *Mashable*, August 2, 2011, http://mashable.com/2011/08/02/newt-gingrich-twitter-followers/.
[10] Rob Dubbin, "The Rise of Twitter Bots," *The New Yorker Blogs*, November 15, 2013.
[11] Anas Qtiesh, "Spam Bots Flooding Twitter to Drown Info About #Syria Protests [Updated]," *Global Voices Advocacy*, April 18, 2011, http://advocacy.globalvoicesonline.org/2011/04/18/spam-bots-flooding-twitter-to-drown-info-about-syria-protests/; Jillian C. York, "Syria's Twitter Spambots," *The Guardian*, April 21, 2011, sec. Comment is free, http://www.theguardian.com/commentisfree/2011/apr/21/syria-twitter-spambots-pro-revolution.
[12] York, "Syria's Twitter Spambots," April 21, 2011.

from -10 fully authoritarian to +10 fully democratic.[13] This preliminary case list suggests that bot usage is often associated with either elections or national security crises. These may be the two most sensitive moments for political actors where the potential stigma of being caught manipulating public opinion is not as serious as the threat of having public opinion turn the wrong way.

Most of the coverage of political bot usage is in new media sources and personal blogs. **Little empirical social or computer science work has been done to understand the wide-ranging creation, use, and effect of computational propaganda.** Existing research is limited to studies on rudimentary bot detection systems, how bots challenge network security, and overviews of bots and botnets—networks composed of bots. Current research fails to develop an understanding of the new political bot phenomena, does not adequately explain the usage of these bots on social media sites, and lacks in any attempt to understand the makers of this technology. Social and computer science research on astroturf campaigns, public opinion manipulation and computational propaganda has been occurring very independently from each other.[14] Social scientists have tended to focus on the organizational culture of engineers and technologists that produce malware, and computer scientists have tended to focus on bot dissemination. The COMPROP team will work together to study both processes and contextualize political bots. While botnets have been actively tracked for several years, their use in political campaigning, crisis management and counter-insurgency is relatively new.[15] Moreover, from the users' perspective it is increasingly difficult to distinguish between content that is generated by a fully automated script, a human, or both.[16]

1 c) Bots and the Internet Census

The first "internet census" was conducted in 2012 by an unknown party.[17] It is not clear that it was a scholarly endeavour but is accepted as a credible study of global botnets. She wrote code that would both count devices and replicate itself so that its copies could help count devices. When activated, it created a botnet that identified 1.3 billion IP addresses used by devices around the world. The author called her script the Carna Bot after the Roman goddess of health. She really did think the exercise was about taking basic measurements of the health of the internet. Her bot worked well, reporting basic information on different kinds of devices, from web cams and consumer routers, to printers and door-security systems. The author of the bot remained anonymous but published her findings as a public service. She exposed two concerning trends with the social application of computing systems. First, she revealed that knowing the default passwords for four pieces of key equipment could give someone access to hundreds of thousands of consumer devices and tens of thousands of industrial devices around the world, from gaming platforms to industrial-control systems. So the world's security experts may be debating the impact of the latest complex hacking attempts from China or the encryption possibilities of quantum computers. Knowing the factory passwords means access to devices once they leave the factory and get connected to the internet.

Second and more concerning, the bot discovered other bots. Carna wasn't the only unauthorized bot checking for open ports on devices around the globe. She exposed several competing botnets, and an enormous, largely sleeping network of bots called the "Aidra botnet" that had compromised as many as 30,000 devices. The bot was designed to hijack not just computers, but gas meters, refrigerators, microwaves, car-management systems and some mobile phones. The bots could attack any network infrastructure for a client with a denial-of-service attack. The author had her Carna Bot perform the public service of temporarily disabling any Aidra bots they found.

[13] Monty G. Marshall and Keith Jaggers, "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2002," 2002, http://www.citeulike.org/group/582/article/369537.
[14] Philip N. Howard, "Introduction: The Hypermedia Campaign," in *New Media Campaigns and the Managed Citizen* (New York: Cambridge University Press, 2006), xxii, 264 p., http://www.loc.gov/catdir/toc/ecip059/2005008088.html.
[15] Kim et al., "On Botnets."
[16] Chu et al., "Who Is Tweeting on Twitter."
[17] Unknown, *Internet Census 2012*, 2012, http://internetcensus2012.bitbucket.org/paper.html.

But the next time someone reboots those infected devices, the bots will be ready to start commandeering devices. Obviously there is a lot of destructive potential behind the malicious botnet exposed, and some might even see her as a threat because she also wrote a script that interfered with network traffic and device function. Understanding the production, dissemination and use of bots and botnets requires tracing the process of engineering decisions, innovations by computer scientists and hackers, the evolutions of social norms of privacy and control, and political values in the use of new technology for propaganda.

Table 1: Political Bot Usage, by Country[18]

Country	Year	Polity	Deployer
Australia	2013	10	State
Azerbaijan	2012	-8	State
Bahrain	2011	-8	State, Outsourced
China	2012	-8	State
Iran	2011	-6	State, Outsourced
Israel	2012	10	State
Mexico	2011	8	Political Parties
Morocco	2011	-6	State, Outsourced
Russia	2011	4	State
Saudi Arabia	2013	-10	State
South Korea	2012	8	State
Syria	2011	-8	State, Outsourced
Tibet	2012	-8	State
UK	2012	10	Candidate
UK	2014	10	Foreign Government
US	2011	10	State, Outsource
Venezuela	2012	2	State

2) Bringing State-of-the-art Social Science to a Contemporary Socio-Technical Problem

Since the bot phenomena is relatively new, there is limited scholarship on how they are produced and little known about their impact on the public sphere. Among the most relevant research is that produced by teams of investigators who explore the political economy of censorship hardware and software. While several researchers study the broad social impact of censorship, there are only a few who are able to provide evidence about both the shared perception that the state is surveilling its public, and specific incidents of censorship that involve disconnections in digital networks.[19] A significant corpus of literature has grown around the use of newer digital media by social movements against authoritarian regimes.[20] While there is a healthy ongoing conversation by scholars on the issue of civil societies' uses of digital media for social and political mobilization, there is little on the use of digital media by political interests operating outside of democracies, seeking to interfere with domestic politics during a regular election. In a sense this is a classic "two-level game" problem in politics: the interactions of political players in domestic politics and those in international politics may once have been easier to distinguish.[21] Today, political communication involves domestic and international actors battling with each other over social media.

---

[18] Torin Peel, "The Coalition's Twitter Fraud and Deception," *Independent Australia*, August 26, 2013, http://www.independentaustralia.net/politics/politics-display/the-coalitions-twitter-fraud-and-deception,5660; Katy Pearce, "Cyberfuckery in Azerbaijan | Katy Pearce," *Adventures in Research*, March 10, 2013, http://www.katypearce.net/cyberfuckery-in-azerbaijan/; Jillian C. York, "Syria's Twitter Spambots," *The Guardian*, April 21, 2011, sec. Comment is free, http://www.theguardian.com/commentisfree/2011/apr/21/syria-twitter-spambots-pro-revolution; Brian Krebs, "Twitter Bots Drown Out Anti-Kremlin Tweets — Krebs on Security," accessed May 14, 2014, http://krebsonsecurity.com/2011/12/twitter-bots-drown-out-anti-kremlin-tweets/; Salley Painter, "Israeli Government Secretly Pays for Pro-Israel Twitter Propaganda," *Top Secret Writers - Conspiracy Theory, Strange Stories and Truth*, accessed May 14, 2014, http://www.topsecretwriters.com/2013/12/israeli-government-secretly-pays-for-pro-israel-twitter-propaganda/; Claudia Herrera Beltran, "En Las Pasadas Elecciones Las Redes Sociales No Fueron Determinantes," *La Jornada:*, March 28, 2012, http://www.jornada.unam.mx/2012/08/28/politica/002n1pol; Freedom House, "Freedom on the Net 2013: Saudi Arabia" (Freedom House, n.d.), accessed April 23, 2014; Choe Sang-hun, "Prosecutors Detail Attempt to Sway South Korean Election," *The New York Times*, November 21, 2013, sec. World / Asia Pacific, http://www.nytimes.com/2013/11/22/world/asia/prosecutors-detail-bid-to-sway-south-korean-election.html; D. Coldewey, "Romney Twitter Account Gets Upsurge in Fake Followers, but from Where," *NBC News*, 2012; Mike Shields, "Bots Infecting Nearly Half of Web Traffic, Per Report," *AdWeek*, accessed May 14, 2014, http://www.adweek.com/news/advertising-branding/bots-infecting-nearly-half-web-traffic-report-152300; insidecroydon, "Jasper Admits to Using Twitter Bots to Drive Election Bid," *Inside Croydon*, accessed May 10, 2014, http://insidecroydon.com/2012/11/26/jasper-admits-to-using-twitter-bots-to-drive-election-bid/.

[19] Ronald J. Deibert et al., eds., *Access Controlled: The Shaping of Power, Rights, and Rule in Cyberspace*, Information Revolution and Global Politics (Cambridge, MA: MIT Press, 2010); Ronald J. Deibert et al., eds., *Access Denied: The Practice and Policy of Global Internet Filtering*, Information Revolution and Global Politics (Cambridge, MA: MIT Press, 2008); Ronald Deibert, *Access Contested Security, Identity, and Resistance in Asian Cyberspace Information Revolution and Global Politics* (Cambridge, MA: MIT Press, 2012), http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=421819.

[20] R. Kelly Garrett, "Protest in an Information Society: A Review of Literature on Social Movements and the New ICTs," *Information, Communication & Society* 9, no. 2 (2006): 202–24; Stephen Marmura, "A Net Advantage? The Internet, Grassroots Activism and American Middle-Eastern Policy," *New Media & Society* 10, no. 2 (2008): 247–71; Katy E. Pearce and Sarah Kendzior, "Networked Authoritarianism and Social Media in Azerbaijan," *Journal of Communication* 62, no. 2 (April 2012): 283–98; Gilad Lotan et al., "The Arab Spring| The Revolutions Were Tweeted: Information Flows during the 2011 Tunisian and Egyptian Revolutions," *International Journal of Communication* 5 (2011).

[21] Robert D. Putnam, "Diplomacy and Domestic Politics: The Logic of Two-Level Games," *International Organization* 42, no. 03 (1988): 427–60, doi:10.1017/S0020818300027697.

2 a) Insight from Research on Political Communication and Social Movements

For civil society actors around the world, digital media and online social networking applications have changed the way in which civic engagement is organized.[22] Social movement leaders from around the world use online applications and digital content systems to organize collective action, activate local protest networks, network with international social movements, and share their political perspective with global media systems.[23] In the past, authoritarian regimes easily controlled broadcast media in times of political crisis; by destroying newsprint supplies, seizing radio and television stations, and blocking phone calls. It is certainly more difficult to control digital media on a regular basis, but there have been occasions in which states have disabled a range of marginal to significant portions of their national information infrastructure.

For several years deliberative democracy researchers have been concerned about the degree to which citizens encounter differing opinions, or are exposed to ideas and public policy options they that might not originate within their community of friends and family.[24] Today political communication systems are significantly more complex and interdependent: social movements project their appeals directly onto the headlines of newspapers in neighbouring countries, and governments shape public opinion formation in neighbouring countries.[25] How should we approach the study of political communication when political actors—not always the state—do more than censor or surveille, but to aggressively push public opinion?

2 b) Insight from Science and Technology Studies of Regime Types

Research in science and technology studies on the role of the state in building and designing public infrastructure may help frame a contemporary study of bots. Civil society is often defined as the self-generating and self-supporting community of people who share a normative order and volunteer to organize political, economic or cultural activities that are independent from the state.[26] And while civil society actors certainly use public information infrastructure they rarely have an active role in engaging with the state on its design. This has led some experts to insist that much contemporary information infrastructure is specifically designed by states to either surveille or capture civic interaction—regardless of regime type.[27] Moreover, many regimes are adopting communication strategies that can only be described as "networked authoritarian": they can be surprisingly responsive to information about grievances that come in over the economic and cultural ties that bind political elites, but still maintain tight control over the flow of information through mandatory points of passage.[28]

While state-of-the-art political communication research may reveal some of the methodological tools for helping us understand the impact of bots, it is regime theory and Science and Technology Studies that gives us the general but substantive questions that need to be answered. If bots and people are embedded in a socio-technical system, what impact have innovations in auto-generated scripts on global social media services had on political discussions and current affairs? Who produces these scripts, or what are the conditions under which innovations in computer science and engineering get repurposed for "computational

---

[22] Bruce Bimber, Andrew J. Flanagin, and Cynthia Stohl, "Reconceptualizing Collective Action in the Contemporary Media Environment," *Communication Theory* 15, no. 4 (2005): 365–88; Brian Still, "Hacking for a Cause," *First Monday* 10, no. 9 (2005), http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1274/1194; Philip N. Howard, *The Digital Origins of Dictatorship and Democracy: Information Technology and Political Islam* (Oxford University Press, 2010).

[23] Jeroen Kloet, "Digitisation and Its Asian Discontents: The Internet, Politics and Hacking in China and Indonesia," *First Monday* 7, no. 9 (2002), http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/1789/1669; Dara N. Byrne, "Public Discourse, Community Concerns, and Civic Engagement: Exploring Black Social Networking Traditions on BlackPlanet.com," *Journal of Computer-Mediated Communication* 13, no. 1 (2007): 319–40; Michelle Shumate and Jon Pike, "Trouble in a Geographically Distributed Virtual Network Organization: Organizing Tensions in Continental Direct Action Network," *Journal of Computer-Mediated Communication* 11, no. 3 (2006): 802–24.

[24] V. Price, J.N. Cappella, and L. Nir, "Does More Disagreement Contribute to More Deliberative Opinion?," *Political Communication* 19, no. 1 (2002): 95–112; Magdalena E. Wojcieszak and Diana C. Mutz, "Do Online Discussion Spaces Facilitate Exposure to Political Disagreement?," *Journal of Communication* 59, no. 1 (March 2009): 40–56.

[25] Krisztina Irion and Giacomo Luchetta, *Online Personal Data Processing and EU Data Protection Reform* (Brussels: Center for European Policy Studies, 2013).

[26] Larry J. Diamond, "Toward Democratic Consolidation," *Journal of Democracy* 5, no. 3 (1994): 4–17.

[27] Katy E. Pearce and Sarah Kendzior, "Networked Authoritarianism and Social Media in Azerbaijan," *Journal of Communication* 62, no. 2 (2012): 283–98, doi:10.1111/j.1460-2466.2012.01633.x; Philip E. Agre, "Surveillance and Capture: Two Models of Privacy," *The Information Society* 10, no. 2 (1994): 101–27, doi:10.1080/01972243.1994.9960162.

[28] Rebecca MacKinnon, *Consent of the Networked: The Worldwide Struggle For Internet Freedom* (Basic Books, 2012); Phillip N. Howard, *Pax Technica: The Impact of Automation on Public Opinion* (New York: Yale University Press, 2015).

propaganda"? Is there a demonstrable impact of bots on news consumption in a networked authoritarian state or in democracies when political actors deploy bots on each other? What is the evolutionary trajectory of this field of computer science, and what are the mechanisms for improving public literacy, generating careful policy oversight, and preventing the abuse of social networking technologies? What does the presence and use of political bots mean for the young people who establish their political identity online and for those in authoritarian countries with restricted news media? How does governmental use of this new propaganda tool effect political organizing efforts and election outcomes? What does such information reveal about algorithmic culture, state interference in digital networks, and digital politics at large? Despite the growing disbursement of this networked disinformation, little is known about the software's creation, dissemination, and capability. Based on the existing state-of-the-art research, COMPROP must investigate three aspects of the role of computational propaganda in European political life between 2015 and 2019: the impact of bots on political discourse; the differences between bot-generated and human posts; and the transitional moment when bot-generated content is accepted and advanced by human users.

Figure 1: Tunisian Politics Blogs, By Keyword

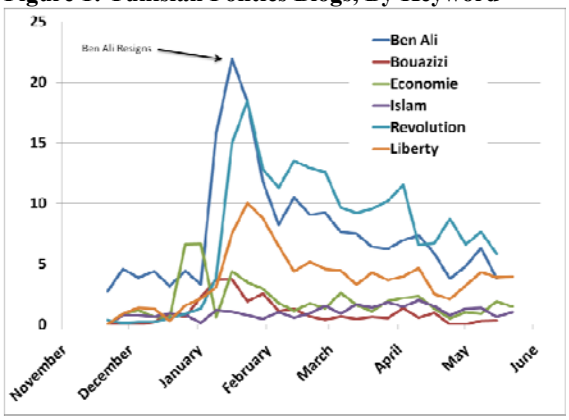
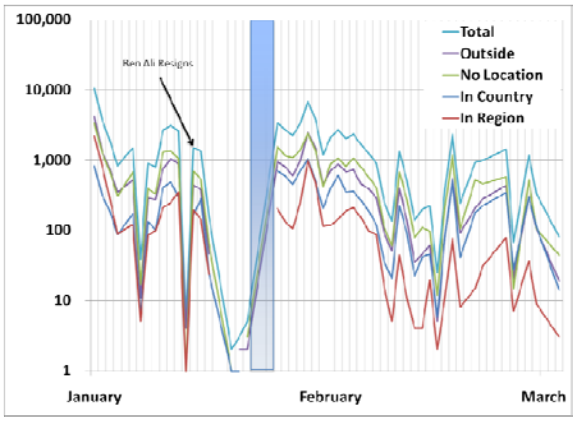


Figure 2: Log N Twitter Users on #sidibouzid, by Location



2 c) Insight from the PI's Previous Research

The PI's previous "big data" established the evidentiary link between social media use, shared grievances, and collective action in Tunisia and Egypt. Analysis of the evolution of both online conversations and offline events allowed for measurement of the degree to which bloggers lead opinion in the public sphere. In December 2011, 5 percent of Tunisian blogs were talking about Bouzizi's self-immolation, Islam, economic conditions, President Ben Ali's leadership, and liberty. By the time Ben Ali was forced to resign on January 14, 20 percent of Tunisian blogs were evaluating his leadership and more than a 1,000 people a day were tweeting about political change. Subsequently, the primary topic for Tunisian blogs was "revolution," until a public rally of at least 100,000 people and the resignation of the old regime's remaining leaders. The day Ben Ali resigned (January 14, 2011), 2,200 people in Algeria, Bahrain, Egypt, Morocco, and Yemen Tweeted frantically about the uprising in Tunisia.[29]

While there is active debate on the weight to give social media among the complex causal factors behind the popular uprisings of the Arab Spring, most analysts now admit it is difficult to tell the story of social change in the Middle East and North Africa (MENA) region without acknowledging the impact of digital networks on the formation of collective action and the cascade effects from country to country. Knowing that such detailed information about public sentiment can be garnered from digital networks of communication between friends and family, what if a political actor wanted to push, promote, or marginalize some sentiment? It is one thing to track such things as a scientist, somewhat after the fact, using big data methods and fieldwork to back up findings. But if bots can intervene and change the terms of debate, or prevent a cascade of dissent, what impact will that have on the public sphere? Knowing that we can model escalations of public outrage helps explain why political actors are now interested in automated ways of promoting keywords or valences in public sentiment.

[29] Philip N. Howard et al., *Opening Closed Regimes: What Was the Role of Social Media During the Arab Spring?* (Seattle, WA: University of Washington, 2011), www.pitpi.org. Blog post data captured beginning November 2010 using the eCairn analytical tool. Tweets captured with Twapperkeeper tool. "Outside Country" refers to Twitter users located outside both the country and the region, and "No location" refers to users with no location data or suspended accounts. The blue bar indicates the period in which journalists began reporting that protests had "thousands" of participants.

3) Research Objectives

Understanding political manipulation in the modern public sphere is not just an interesting research or theoretical challenge, it is an important normative initiative that the combination of social and computational science can undertake. In view of the state-of-the-art within political communication and socio-computational systems, COMPROP can lead the research agenda through five objectives:

1) To study the way bot makers and bot detectors work, think, learn, design, and interact by both mapping professional innovation networks and ethnographically experiencing lab and business environments.
2) To provide new knowledge of the way bots scripts operating over important social networking applications like Twitter and Facebook work, not only in terms of how they use devices and replicate themselves, but in terms of the content they spread and real-world influence or impact they have.
3) To develop innovative models through mixed-method qualitative, comparative and quantitative research methodologies, models for bot deployment on specific political issues or during important elections.
4) To demonstrate the impact of political bots and shape a policy agenda for social media firms, civil society groups and government policy makers on the deployment of political bots.
5) To validate the models on bot deployment through actionable scenarios for how to improve bot detection by crowd-sourcing bot-detection during critical elections in the later years of the project.

2. b) Methodology

While the scholarship on computational manipulation in political communication is almost non-existent, the recent scholarship on the broad impacts of technology diffusion on politics does help us form research questions and provides some sensible ways to proceed methodologically. Drawing from multiple sources, it is possible to do a comparative analysis of the myriad incidents in which government officials decide to censor their online publics. By collecting as many known incidents of state intervention in information networks, we can map out the contours of crisis situations, political risks, and civic innovations to understand the new intersections between computational interventions and political impact.[30]

1. Specific Research Questions

Through the construction and analysis of a globally comparative event dataset of political bot usage the proposed COMPROP project engages with the following research questions:

Research Questions for Work Package 1. The first work package involves international qualitative fieldwork, and is driven by the questions—who makes bots, why, and under what circumstances? Building what we know from existing research, the PI hypothesizes that bot activity will peak during major elections in Europe over the next few years, and on average there are between 5 and 7 national elections each year. There are also Europe-wide elections and occasional referenda that can be quite controversial. The best available calendar estimates of elections only cover 2015-2018. It is safe to assume that there will be other political crises over this time period and other elections in 2019 in which bot activity should be tracked. COMPROP's research timeline can plan with the electoral calendar in mind, and be flexible and adaptable in its study periods to be able to do data collection as needed.

Research Questions for Work Package 2a and 2b. The second work package involves the creation of an original event data set of bot deployment—what is the impact of bots on political discourse? Knowing what their impact on public opinion or information networks, what explains variations in their impact? The case list of known deployments of political bots will grow rapidly over time. Indeed, as new scripts are written and demonstrate effectiveness, it is likely that more and more political actors will attempt to use digital media and social networks for manipulation and control. Answering these questions requires a crafted combination of qualitative, comparative and quantitative data, and a methodological toolkit that can both serve the needs of traditional statisticians working in the social sciences and render new knowledge through cutting edge fs/qca statistics.

[30] Matthew Carrieri, Ronald Deibert, and Saad Omar Khan, "Information Infrastructure and Anti-Regime Protests in Iran and Tunisia," in *State Power 2.0: Authoritarian Entrenchment and Political Engagement Worldwide*, ed. Philip N. Howard and Muzammil M. Hussain (Farnham, UK: Ashgate, 2014), 45–56; Jennifer Earl et al., "The Use of Newspaper Data in the Study of Collective Action," *Annual Review of Sociology* 30, no. 1 (2004): 65–80, doi:10.1146/annurev.soc.30.012703.110603.

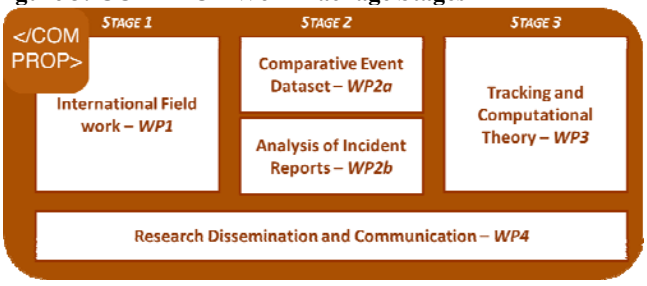
Research Questions for Work Package 3. The third work package is forward looking. Can we model the immergence of political bots? After the initial period of fieldwork and analysis of the first few years of events in the database—and in consultation with the community of experts built around the **"COMPROP Budapest Bot Workshops"**—we will be able to theorize a system of how to crowdsource and crowdseed the nomination of hashtags and bot accounts for live tracking by simple scripts that we write and test. Being able to track bot activity and impact in real time will be of immense value not only to the engineers and managers of infrastructure and digital services in Europe, but will contribute to policy maker's awareness of problems and public awareness of digital campaign tricks.

COMPROP will answer these research questions and achieve its research objectives through qualitative fieldwork among the people who create bots. Learning about their communities, commissions and techniques will help us better understand the design process—in terms of strengths and weaknesses. Fieldwork will also allow us to begin building an event log of incidents where bots were commissioned, designed, and deployed. Additional research among news sources and specialized programming sources will allow the development of a larger database of cases that can be analysed for the causal narrative that may connect particular political contexts and issues with bot design and implementation and political or opinion changes. We will develop an original computational model for how to crowd source the detection of political bots, a system that might work in conjunction with Natural Language Processing to help identify not just spam-driven traffic spikes but substantive content that is politically motivated and bot-driven. Detecting bots is one of the most computationally difficult tasks and often requires human judgement. But behind the bots are norms and innovation networks, and by mapping them out we can get closer to understanding their modes of operation and the trajectory of design options for bot builders. This plan is mapped across four distinct work packages.

2. Research Plan of Work Packages

To answer these questions, the PI proposes a three stage research design (Figure 3) in which a dedicated research team builds from grounded knowledge, qualitatively gathered, through an original comparative event data set, to theoretical concepts that advance our understanding of how to track bot activity both socially and computationally. In part, the selection of field sites will be driven by the course of international crises over the next 5 years. The PI hypothesizes that elections in particular are likely to involve the production of politicized social media bots. Internationally, likely sites for such activity in the next year include the Syrian presidential elections in June 2014, the Turkish Presidential elections in August 2014, and the Brazilian Legislative elections in October 2014. In the European context, parliamentary systems make it difficult to predict with certainty the timing of electoral contests, and other kinds of political, economic, and security crises will create additional moments for study for which the team must be ready. Nevertheless, Table 2 reveals the most likely moments over the next 5 years when political bot activity will peak in Europe and Europe's neighbouring countries. These will be prime moments for COMPROP data gathering efforts. Tracking bot production during these sensitive moments—and being responsive to other international crises as they develop—will help create the snowball sampling technique on bot producers.

Figure 3: COMPROP Work Package Stages



2.1. Work Package 1: International Fieldwork

The first and early stage of the project involves rigorous social science fieldwork methods to gather knowledge of how bot designers operate professionally, both in terms of working as an innovative network of engineers and as a professional network competing for clients in a market for computing services. PI Howard has a demonstrated record of working with hackers, hacktivists, spammers, and political campaign managers whose work violates most people's privacy norms and technology values. There are three target groups for the interviews:

(i) ***Makers of Political Bots***. These individuals and firms are internationally distributed, and some of the engineers behind political bots actually work at major advertising firms. The process of making contact and confirming willingness to participate in the interviews has already begun.

(ii) ***In-House Engineers***. All of the social media firms with high profile social media services employ computer engineers to detect bots. The process usually involves some algorithmic identification of problematic accounts that publish too quickly, but with the growing sophistication of bots human confirmation is often needed before accounts are sanctioned or deleted.

(iii) ***Industry Research Computer Scientists***. There are a few third party organizations dedicated to identifying bots. A few work for online services such as Status People and Truthy, but there are small but leading research teams at Microsoft Research and the University of Illinois-Urbana Champaign that the PI has easy access to.

With a team of post-docs and advanced graduate students, the PI will interview bot makers who produce the automated scripts for political actors. Contacts within the first two groups have already been made. The PI has already begun grooming network ties to the makers of political bots and the in-house engineers. Several are based in Russia, Bahrain, Hungary, and other countries within Eastern Europe and MENA. PI Howard has begun contact overtures, several bot makers have already agreed to be interviewed, and the PI has a demonstrated record of being able to work ethnographically with the hacker community. Since one of the broad impacts of this project will be to improve the efficiency of bot detection and raise the shared understanding of how bot detection systems should evolve, and since the PI has his own professional ties to university-based researchers, contact with the third group will be made early in the implementation of this project.

.

2.2. Work Package 2a: Building Comparative Event Data Set

The second stage of the project involves the construction of an original event dataset more comprehensive than any previously collected. Event datasets have become particularly powerful tools for understanding trends in socio-computational systems, including global digital activism, compromised personal records, and government interference with digital switches.[31]

To create this data set, a group of trained and supervised graduate student coders will review news stories created by both citizen and professional journalists which describe the impact of bots on political discourse. The PI has identified several dozen cases through his exploratory research (many identified in Table 1 above), and will use a purposive and snowball sampling technique to identify cases. Research assistants will read each source and assign values for qualitative and quantitative variables defined in a crafted codebook. The perspective will be global, and the objective in this research is to build a typology of the evolving use and impact of what the PI is calling "computational propaganda" in both democracies and authoritarian regimes around the world. This part of the project will involve the PI, postdocs, and graduate students, and will render important insight into the size of the consulting industry that produces political bots, and identify the key network actors the team needs to contact. Case coding will follow the high standards of these datasets and involve a small team of specially trained coders who participate in training, collect cases, research details, enter case information, participate in retraining, and get evaluated through inter-coder reliability scores.

Following the classical methodology of the study of unusual phenomena in technology diffusion and usability, the team will begin sampling bot use by means of news reports about them.[32] The media-based approach to data collection is especially valuable when the phenomenon at hand is particularly new. The method for analysing these texts will be a content analysis, a systematic means of textual analysis which endeavours to have all observers come to the same conclusions about the content of the text. This inter-coder agreement increases the reliability and also the authority of the attendant findings.[33] Since a core COMPROP hypothesis is that elections will be likely moments of bot deployment, Table 2 identifies the countries most likely to have impacted infrastructure and public discourse.

Unlike many content analyses, in which unit of analysis and unit of observation are one in the same, in this study the two will be different. The unit of analysis is the bot campaign while the unit of observation is the news report about that campaign. This means that this stage of the analysis involves the indirect study of computational propaganda and gives us the added value of using third-party sources that can be evaluated for

[31] Edwards, Howard, and Joyce, "Digital Activism and Non-Violent Conflict"; Kris Erickson and Philip N. Howard, "A Case of Mistaken Identity? News Accounts of Hacker, Consumer, and Organizational Responsibility for Compromised Digital Records," *Journal of Computer-Mediated Communication* 12, no. 4 (2007): 1229–47.
[32] Earl et al., "The Use of Newspaper Data in the Study of Collective Action."
[33] Klaus Krippendorff, *Content Analysis: An Introduction to Its Methodology* (Sage, 2012).

trustworthiness [34]. Media bias is a legitimate concern that can be mitigated by relying on a variety of news outlets and on amateur as well as professional sources. This method alone will not completely nullify media bias, but combining the sampling strategy with knowledge garnered by international fieldwork with bot makers and bot detectors will help significantly. While traditional news sources and peer-reviewed journal articles will provide some entries, the PI expects that the interviewing process will drive the snowball sampling strategy to such a degree that the team will come close to gathering the universe of cases in which any political-motivated bot provided was designed and released on social media.

Table 2: Hypothesized Political Bot Deployment in Europe, 2015-2018

Country	Election Type	Year
Estonia	Parliamentary	2015
Finland	Parliamentary	2015
Poland	Presidential	2015
Slovenia	Parliamentary	2015
Spain	General	2015
Turkey	General	2015
UK	General	2015
Georgia	Parliamentary	2016
Lithuania	Parliamentary	2016
Romania	Local, Legislative	2016
Scotland	Parliamentary	2016
Slovakia	Parliamentary	2016
Armenia	Parliamentary	2017
Bulgaria	Parliamentary	2017
French	Presidential	2017
Germany	Parliamentary	2017
Iceland	Parliamentary	2017
Norway	Parliamentary	2017
Serbia	Presidential	2017
Armenia	Presidential	2018
Georgia	Presidential	2018
Czech Rep.	Presidential	2018
Cyprus	Presidential	2018
Finland	Presidential	2018
Russia	Presidential	2018

2.3. Work Package 2b: Analysis Comparative Event Data Set

The population of cases will include all incidents in which bots, botnets, or automated scripts are known to have had an impact on political life. Thus, their impact must have been reported or felt in the news media, or by a user group, and to be captured in this analytical frame the code must be more than spam or have had an impact on traffic. For it to be included it must have an observable impact on some political conversation. International fieldwork will not only help map out the network of operators making and disseminating bots, but will also help in the collection of cases.

Key variables will include the attributes of the bots (type of platform, author, complexity, programming attributes), attributes of deployment (deployment date, speed of deployment, number of devices, users, or user accounts impacted), network impact (devices disabled, services disabled, traffic impact), messaging (issue domain, event driven, positive or negative valence, rhetorical techniques, message codes, images, image codes) and political outcomes (for attacker, for target, for involved users, for level or volume of discourse). While specific models are difficult to estimate given the paucity of relevant literature and current lack of fieldwork, fuzzy logic models are likely to be productive.

Fuzzy Logic Statistical Modelling of Impact. There have been a few singular case studies of bots that were developed for political applications, but new cases are being detected with alarming frequency. The comparative perspective taken in this investigation will be defined by incidents of bot driven, enhanced, or enabled regime change. Methodologically, the comparative approach is powerful and productive in that it confronts theory with data. Sometimes this approach is called "set-theoretic" in that attention is given to consistent similarities or differences across a set of cases, especially the causally relevant commonalities uniformly present in a given set of cases.[35] Set-theoretic reasoning allows for fine gradations in the degree of membership in the set of outcomes—in this study the deployment of a bot and some political impact.

Fuzzy logic is most appropriate because of the interesting mixture of variables concerning political context and bot design. Bots are often composed from scripts that have been used before, and it is their deployment strategy that involve the most creative ingenuity. The impact of bots probably varies with whether or not it is a foreign government behind the commissioning, or a political party. The impact certainly varies with design attributes, language, and country context. It almost certainly varies by which platform is being used. Moreover, the relationship between bought design, social acceptance, and political impact will be evidenced with both the successes and failures. Fuzzy logic allows for many different kinds of fixed crisp attributes and fuzzy categories and makes use of all real-world cases, successful or otherwise.

2.4. Work Package 3: Tracking and Modeling Botnet Emergence

There are disparate projects dedicated to locating and detecting bots. The interviews with bot creators will help the team gain deep insight into three critical aspects of bot generation: a) What are the methods they use to influence online conversations? b) What hashtags/events do they target? c) How do they make bot posts

[34] William Dutton et al., *The Internet Trust Bubble: Global Values, Beliefs and Practices* (Oxford, UK: World Economic Forum, 2014).

[35] Benoît Rihoux and Charles C. Ragin, *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques* (Sage Publications, Inc, 2008).

similar to human-like posts? The interviews with the bot trackers – or the digital detectives as they like to be named—will help us better understand the methods used to track bots and learn how to control the domains that bot creators use to infect computers. After gaining insights from the two different groups of experts the plan is to start prototyping new bots using the same techniques. The embedded knowledge the team will gain from creating and disseminating bots will enable us to produce insightful recommendations on how to detect social bots, protect free speech and promote a healthy digital public sphere.

In this third stage of the work, conducted after fieldwork and the construction of the event dataset, the team will seek to elaborate and extend existing models of bot detection. This third piece of research will seek to account for the qualitative findings from the prior two components to model when computational propaganda will be used, and test the model by tracking and targeting Twitter and Facebook bots through targeted datasets that are assembled around politically critical events. While much of the innovation here will be driven by the findings in the first and second stage of this COMPROP project, it is likely that the PI will take advantage of context-relevant crowd sourced knowledge, existing scripts that he has for Twitter analysis, and the additional computing skills of consultants who are experts in other social networking applications. Moreover, while several researchers have made use of a variety of "crowd-sourcing" methods for research purposes, COMPROP will employ an innovative "crowd-seeding" strategy that both welcomes user input on which bots to track but also ensures a systematic distribution of mandatory reporters across countries.[36] Constructing an online interface to have users nominate hashtags and bot accounts worth tracking may allow us to efficiently apply crawling and capturing applications and investigate events for possible inclusion in the event data set.

2.5. Work Package 4: Research Dissemination and Communication

Central to COMPROP's impact strategy is to bridge gaps across the social and computer sciences, and connect the knowledge deliverables to public conversations about trust in the future internet. COMPROP is committed to in-progress publishing and open access to the knowledge generation.

Social media are key to COMPROP's dissemination. The PI's Twitter account has over a thousand followers, many of whom are other important nodes in research networks (and are not bots!). The Oxford Internet Institute, has its own strong web presence. COMPROP will have a dedicated website and online platform for distributing data, and the release of original datasets—on an annual basis—are actually important benchmarks for the project. COMPROP's digital strategy will allow broad access to links, resources, examples, case studies, and raw data. The PI has integrated project **blogs** into a cutting edge version of the ethnographic method that involves writing up basic field for publication to the team. This can serve as an early alert system for the emergence of new bots, and functionally helps the ethnographers begin to transition their observations from loose field notes to research memos. Contemporary researchers also do a significant amount of cross-posting, and the PI's connections will allow COMPROP posts to be shared with Princeton's highly trafficked *Freedom To Tinker* blog, and Columbia University's *Tow Center For Digital Journalism* blog. The PI has also been just invited to blog for the LSE's *Media Policy Blog*. This means active dissemination of results to scientific/academic communities, as well as to other relevant stakeholders (policy makers, technology industry, start-ups), journalists and the interested public.

Timed shortly after the annual release of data on bot activity, the **COMPROP Budapest Bot Workshop** will be held and researchers from the social and computer sciences, policy makers, and industry partners will be invited to Hungary for a two day event. Frank discussions about the evolution and use of political bots will allow for a sharing of knowledge, fill out the case list, and allow COMPROP's specific research strategy to stay current with the evolution of bots and political events on the ground. Not only will these briefing and workshop sessions bring the international expert community to Budapest, but they will strengthen network ties within European teams of bot tracking teams in science and industry. These workshops will be project milestones that allow us to engage others with the raw data, and to release the latest contributions to the policy paper series from that year. Because the PI is also a faculty member at the OII, this research will

[36] Judd Antin and Aaron Shaw, "Social Desirability Bias and Self-Reports of Motivation: A Study of Amazon Mechanical Turk in the US and India," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12 (New York, NY, USA: ACM, 2012), 2925–34, doi:10.1145/2207676.2208699; Aniket Kittur et al., "The Future of Crowd Work," in *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, CSCW '13 (New York, NY, USA: ACM, 2013), 1301–18, doi:10.1145/2441776.2441923; Peter van der Windt, "From Crowdsourcing to Crowdseeding: The Cutting Edge of Empowerment?," in *Bits and Atoms*, ed. Steven Livingston and Gregor Walter-Drop (Oxford University Press, 2014), 144–55, http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199941599.001.0001/acprof-9780199941599-chapter-10; Shirky, *Here Comes Everybody*.

actively involve and employ graduate students at Oxford. The searches for postdocs will be open broadly, but locally engaged research teams will be used for coding cases, planning workshops, and developing research deliverables. PI Howard has a strong record of involving students in original research, and COMPROP will be another opportunity to mentor the next generation of information policy makers.

2.6. Gantt Quarterly Workflow

The Gantt Table (see below) maps out the quarterly flow of the four work packages with tasks and roles from 2015 to 2020. The fieldwork will be a staged project activity. Event dataset design will occur after getting into the interviews because through those interviews the team can begin to collect the case list for the event dataset. Team conversations about tracking and models for detecting bots can begin as soon as the team has some substantive field notes to drive team dialogue forward. The mentoring activities will occur at every stage of the project, and engagement with scholarly conferences can begin mid-way through the project when the team has enough interview and event data to be able to report results.

3. Expected Impacts

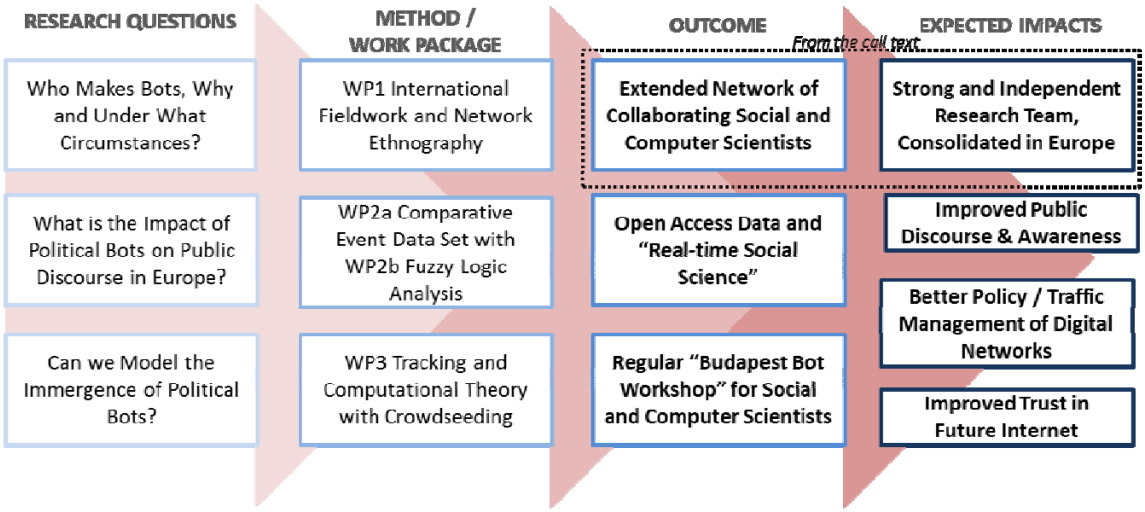
The event dataset produced by this research will have a broad impact on the network of industry and university researchers working on the problem of computational propaganda. Indeed, while our COMPROP proposal will result in new computational theory about bot tracking based on our grounded study of bot producer networks, the PI is certain that the dataset will have a broad impact on the network of researchers working on detection. The cleaned dataset and codebook will be specifically shared with the teams of bot detectors who participate on our study.

Our team is after specific evidence of how learning, design and repurposing occurs among bot makers, not simply an archive of bot features. Timely research on the bot activity can best serve European foreign policy experts, computer scientists—and democracy—now. Policy makers in Europe need greater literacy in the impact of innovations in science and technology on politics, beyond what pundits provide. Social science research on human computer interaction is now extremely relevant for public policy. COMPROP will generate new knowledge for the benefit of:

- Europe's technology industry, which can integrate findings to improve digital traffic management;
- Civil society, which will understand the emerging impact of digital manipulation over hardware and software and develop sophistication with tracking and response as needed;
- Policy makers, who will have more tools for understanding the balance between free speech, political manipulation, policy oversight of media and elections;
- Scholars, who will have new methods of integrating qualitative and computational research, and deeper appreciation of the value of working between social and computer science.

Figure 4 maps out the path from research questions to method choice and work package, and summarizes both the key project outcomes and the expected impacts. The **scholarly output** of scientific papers, monographs, conference papers and presentations, coupled with the regular publication of the COMPROP policy paper series, dedicated project website and blog posts will ensure wide dissemination of our findings.

Figure 4: COMPROP, From Research Questions To Methods, and Outcomes and Impacts



Moreover, our social media outreach and project specific dissemination events (annual **COMPROP Budapest Bot Workshop**) will help anchor the new interdisciplinary community built around our work.

The outcomes of this project will be better awareness of the impact of bots on political life and better sophistication at identifying the content in our social media stream that is generated by automated scripts. By the end of this project the PI and research team will have raised public awareness of the impact of bots on political life and the average level of sophistication at distinguishing bot-generated content in social media feeds. The policy papers, original research and dissemination we do to journalists will raise the issue in news coverage and improve the ability of journalists to cover political manipulation over social media networks as a public interest news issue. Engineers and computer scientists will have a better sense of the targets of their efforts at blocking bots and preventing them from eating up bandwidth, clogging traffic and polluting social interaction. Policy makers across Europe who specialize in the oversight of free and fair elections, media regulation and freedom of expression will be better equipped to understand how bots impact their policy domains and find public policy tools for discouraging their use and minimizing their impact.

4. PI Leadership, Capacity and Creativity

PI Howard has recently moved to Oxford University from Central European University, and this COMPROP project would provide the support needed to advance and consolidate his research agenda on Europe.

4.1. Success of Prior Projects

A key reason for the success rate of creative proposals to funders is his capacity to follow up research findings with a two-part dissemination strategy. First, my scholarly output is largely driven by the creation of original event data sets that are openly shared immediately after publication. This research process has resulted in unique contributions particularly in science and technology studies, with studies of hackers and privacy norms, digital activism, and the causes and consequences of the Arab Spring. This makes innovative frame-changing datasets immediately available to research colleagues. Second, the PI always designs an extended dissemination strategy for new research findings. This usually involves translating scientific conclusions into accessible texts for journalists and public policy makers—ultimately making findings available to the broader public

4.2. Fit with Current Research Trajectory

a) The project will continue PI Howard's research trajector in political communication and socio-computational systems. COMPROP will provide evidence for PI Howard's forthcoming book with Yale University Press: *The Internet of Things: Pax Technica and the Coming Challenges to Democracy*. This is already contracted, which means the PI already has a prominent dissemination opportunity ready for COMPROP outcomes. Building on field work in nine authoritian regimes and developing countries, this manuscript will make the argument that European policy makers need to consider how new technologies and technology cultures—especially those around new networked and sensing devices—are may be degrading open political conversations. The "pax technica" is the stable arrangment between industry and government that makes technology access a public and industry priority but means that some of the actual practices of building and using new information infrastructure have introduced unregulated opportunities for political manipulation, mostly through bots and digital rights management over the internet of things. He argues that the "pax technica" that is emerging in Europe and the U.S. is integrating the political cultures of information technology communities that support open data, open networks, and open societies into notions of governance and providing a strong alternative to the regimes of authoritarianism and control emerging around these technologies in less democratic regimes. The project also extends the "network ethnography" methods pioneered by Howard for linking rigorous qualitative data collection and case selection with social network analysis.[37]

5. Conclusion: COMPROP as a Field Defining Project

Behind the bots are norms and innovation networks. By mapping out innovation networks we can get closer to understanding modes of operation and design trajectories for bot builders. The PI has a demonstrated record of scholarship involving socio-technical systems of hackers, computer scientists, and innovation networks. Moreover, work packages involve active engagement with teams of computer scientists involved

[37] Philip N. Howard, "Network Ethnography and the Hypermedia Organization: New Media, New Organizations, New Methods," *New Media & Society* 4, no. 4 (2002): 550–74.

in bot detection, the fieldwork will be with the authors of automated scripts, and the budget specifically provides post-doc opportunities for computer scientists interested in the political impact of bots.

5.1. Involvement of Research Subjects

PI Howard has a demonstrated record involving deviant computer professionals and engineers in research. The PI has already located several potential research subjects (contact with them will be resumed with Human Subjects approval has been awarded). Involving both producers and the teams of computer professionals who work in industry, government, and the technology industry means being able to say something about the full production cycle and the big picture impact of bots and algorithms on political culture. In other words, this project will be made even more feasible by working with the broad professional community of computer scientists and relevant policy regulators working on the problems caused by bots. In the end this **research will benefit both industry and the public sector**, by facilitating interaction between the corporate engineers working for social media firms in Europe, the research staff of industry labs who are dedicated to improving bot detection, and the policy makers who need to learn about the impact of political bots on public trust.

5.2. Evidence of Creative and Independent Thinking

The PI's first single authored book, *New Media Campaigns and the Managed Citizen* (Cambridge University Press, 2005) applied the new method of "network ethnography" to the study of election campaigning in the United States, and provided the first large demonstration of how technology diffusion had allowed more citizens access to more information but at a cost of privacy and personal control. This is now accepted as received wisdom, and subsequent research has both built on and confirmed the PI's findings. The fieldwork for this book was conducted as a grad student under the supervision of Professor Charles Ragin, and was done during the 2000 Presidential election in the United States. However, Cambridge University Press reviewed the book manuscript and asked for additional archival work on the 1992 and 1996 elections, and asked for additional fieldwork on the 2004 elections. In its final form as a published book, most of the book had been rewritten after completing his doctoral work—independent of his PhD supervisor Dr. Charles Ragin—and while working as a junior faculty member in a tenure track appointment.

The PI's second single authored book, *Digital Origins of Dictatorship and Democracy* (Oxford University Press, 2010) developed fuzzy set statistical models on the impact of technology diffusion on political Islam in 75 countries. This study argued that there were significant changes underway in collective identity, gender politics, and the interpretation of Islam, changes that were in part due—in a complex causal recipe—to the rapid diffusion of social media. It is a creative and substantive application of the new toolkit that develops deliberative democracy theory in significant ways. Most important, the conclusions foreshadowed the issues that arose in the Arab Spring, and the PI become globally recognized for this prescient work with awards and fellowships and the world's leading research institutions including Harvard and Princeton University (the PI declined the Harvard University appointment, accepted the Princeton University appointment).

6. Justification for ERC Support

It is very likely that foreign governments and political actors are already using bots to manipulate public opinion in Europe. This project will greatly expand our understanding of how this is done, and advance the conversation among researchers in the computing, engineering, and social scientists about the size of the problem and the possible solutions. But doing this well means building a multi-disciplinary team under the leadership of someone with experience working on the social impact of innovations in computer science.

The team will interview bot-makers in key countries in order to generate insight into the inner workings of political bots and botnets. This information will allow keen understandings of the nuances of hashtag usage, back-end social media site construction, and bot mechanics. COMPROP will demonstrate how bots impact the social systems in which they are deployed and how specific aspects of computational propaganda, data used for coercion, discrimination, and control, play out globally. Such research will potentially generate new theoretical understandings and will uniquely contribute to the standing theory of a variety of fields associated with the computer and social sciences.

Gantt Table: Quarterly Workflow For Work-packages (WPs), with Tasks and Roles from 2015-2020

	Start-End Quarters	2015 Winter	2015 Spring	2015 Summer	2015 Fall	2016 Winter	2016 Spring	2016 Summer	2016 Fall	2017 Winter	2017 Spring	2017 Summer	2017 Fall	2018 Winter	2018 Spring	2018 Summer	2018 Fall	2019 Winter	2019 Spring	2019 Summer	2019 Fall	2020 Winter	2020 Spring
Project Quarter		1	2	3		4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Prior to Award																							
b) Finalize approval from human subjects committee	..																						
c) Plan and prioritize network contacts	..																						
WP 1: International Fieldwork																							
a) Background research on bot makers and detector labs	1-2																						
b) Develop interview script and observation plan	1-3																						
c) Conduct fieldwork	2-7, 14-19																						
d) Exit fieldwork: observational memos, summary surveys	7-9, 18-20										●											●	
WP 2a: Building Comparative Event Data Set																							
a) Develop coding instrument, train coders, run pretest	1-4					○																	
b) Snowball sample with cases from fieldwork & Bot Workshop	4-29																						
c) Code incidents, periodic intercoder reliability tests	4-29																						
d) Cleaning and releasing; variable based reliability tests	4, 8, 12, 16, 20					●				●				●				●				●	
WP 2b: Analyzing Comparative Event Data Sets																							
a) Descriptive statistics for trend analysis	4-5, 8-9, 12-13, 16-17, 20																						
b) Traditional statistical modeling	5-6, 9-10, 13-14, 17-18, 20																						
c) Fuzzy logic modeling	6-7, 10-11, 14-15, 18-20							○				○				○						○	
d) Use model findings to revise and improve codebook	7, 11, 15																						
WP 3: Tracking and Computational Theory																							
a) Review fieldwork and event data for trends	10-13																						
b) Develop socio-technical model for deploying bot detector	1-4, 13-16																						
c) Test crowdseeding system during major European elections																							
d) Deploy and test prototype bot detector	4-7, 16-20								○													○	
WP 4: Research Dissemination																							
a) Conference presentations, Budapest Bot Workshop	7-20						●				●				●				●			●	
b) Article submissions	10-13, 19-20+																						
c) Release event data set	4, 8, 12, 16, 20					●				●				●				●				●	
Post Award																							
a) Follow on analysis in scholarly papers	..																						
b) Project closing documents and filings	..																						
c) Additional dissemination activities	..																						

Milestone = ● Deliverable = ○ Milestone and Deliverable = ●

2. c) Resources

Project Costs

Percent of working time the PI dedicates to the project over the period of the grant:	**50%**

Budget Justification and Project Objectives

Personnel
PI: Support is sought for 50% of the PI Howard's salary including the benefits to which Oxford employees are entitled.

Postdocs: Support is sought for two postdocs. At any given time during the life of the project one postdoc should be for a social scientist and the other for a computer scientist. A core objective of this project is to improve the literacy of the social sciences with computational propaganda, and another is to make the computer sciences more aware of the political impact of innovation in automated scripts for content control over social media. Thus, one postdoc would go to a social scientist having completed doctoral work on some aspect on the political impact and policy oversight of algorithmic culture, bots, or other relevant topics. The other would go to a computer scientist with a relevant background who can help with fieldwork, disseminating project findings with the wider community of computer sciences, and lead the theoretical and design conversations about how to detect political bots. The project will have 40 months of support from a social science PostDoc and another 40 months of support from a computer science PostDoc over the life of the grant.

Other personnel: This amount includes four other types of personnel. 1) Hourly graduate student research assistance. Involving students in original research is another objective of this project. The Oxford Internet Institute has a talented pool of students who can be involved in various aspects of data case collection, analysis, and research dissemination. COMPROP will have 25 months of support from graduate student researchers at Oxford University over the life of the grant.

All personnel costs are added employers social security contributions at a rate applicable at the time of application in the United Kingdom.

All staff on the grant will be employed directly by the University of Oxford, including the PI who will be employed for the full duration of the grant.

Travel: The international fieldwork component of this project is important because it will provide the grounded and ethnographic observations of how the labs of bot makers and bot detectors work. Most of the bots that are used on European publics are designed and launched from Russia, the United States, and a few other countries. Field sites have been tentatively identified in Moscow, Bulgaria, London, San Francisco and Bahrain—these are the places where known teams of bot makers and bot detectors work. Teams of two-ideally one computer scientist and one social scientist, will conduct team interviews. Additional travel to hacker conferences will be needed, and in part the travel schedule will be set by the appearance of new innovations in automated politically manipulative scripts. Involving the postdocs and graduate students in the fieldwork is an important mentoring goal, so a significant portion of the travel budget can be used to support their experience with international fieldwork. The timing for destinations will on which consultants and bot producers are active at a critical time. Prime conferences include Computer Supported Collaborative Work, Theorizing the Web, International Communications Association and the iConference. Such conferences tend to have attendance fees and the fees will be met from the project budget, and additional small stipends for conference participation may be provided by the project budget.

All travel costs will be claimed in accordance with Oxford University's normal institutional management and accounting practices. The travel costs of any experts travelling for the purposes of the grant are included in the budget.

Equipment: To succeed, this project requires some modest but especially dedicated equipment above what is normally provided by the university. First, the project requires a dedicated server infrastructure, which includes a server that can handle the project website and data files, and operate as a secure service for project personnel. There is a slim chance that in studying politically motivated bot activity the project will become a candidate for DDOS attack. Support for modern secure server, with peripherals and cooling rack. In addition, two ruggedized laptops are needed for project personnel, as a significant amount of the qualitative fieldwork is going to require travel and dependable laptops, which would be used for solely research purposes and can handle large databases.

While the University provides IT support and computer network facilities it is normal institutional practice for research projects to fund all researcher requirements for computer hardware. Costs are identifiable and directly attributable to research projects in the University's accounting system. Computers are claimed under consumables and not as equipment as it is Oxford University standard accounting policy to not capitalise items under £25,000 in value: they are written off in the year of purchase

Other Goods and Services: Consumables: Support for specialized cloud applications, additional software, services for disseminating research findings over social media.

Other Goods and Services: Sub-Contractors: 1) Writing stipends for external experts to contribute to working paper series. This project will design a high profile policy paper series that features a mixture of policy analysts and computer scientists writing about the critical impact of information infrastructure on policy debates. Inducements will be offered to help secure high profile authors and help ensure the timely delivery of manuscripts written at an accessible level for a broad public rather than a specialized computer and social science public. 2) Designers / copyeditors for policy paper series. Support is sought for professional design and copyediting services—beyond what is normally provided within the university—for the policy paper series. 3) Computer programmers. As the project nears the stage of being able to theorize about ways to catch political bots operating "in the wild", additional support from a computer programmer will be needed. In the early years of the project this will involve exploratory work and support of the postdoc, and in later years this will involve testing and refining any application designs that may develop special features or be platform specific to the Facebook or Twitter API. No staff costs will be charged to the grant for project collaborators or external experts.

These costs are determined according to the usual accounting and management principles and practices of the organisation; they will be appropriately substantiated and directly linked to the project, adequately recorded, identifiable and verifiable.

Publications: Support is sought to ensure that the scholarly output of this project is available for open access, and the primary item here is the printing of hardcopy version of project research output and open access fees. Publication support will be used to ensure that the findings of these projects are openly accessible. Open access is especially important for this project because the findings will be relevant for both the computer and social sciences. Any raw data about how bots work can be used to improve detection, raising the chances that the manipulation of political discourse can be stopped early and the chances that information scientists in research or industry can catch and block bots quickly.

These costs will be used for the sole purpose of disseminating team research.

Other: Support is sought for auditing services.

C1 – Subcontracting Costs: Subcontracts will be needed to guarantee access to the Twitter and Facebook "firehose" of data. Pricing varies by service needs, native hosting capacity and the infrastructure/skill set of the analysts, but this figure is based on the current estimates of costs as of May 2014.

The PI is committed to spending 50% of his time advancing this research agenda and managing the research effort. Moreover, the PI has a demonstrated record of involving students in original research and in integrating research questions into the pedagogy occurring in the classroom. This means that while 50% of his time will be formally released for teaching, the PI expects to integrate the research agenda into the

remaining formal pedagogy in the classroom and to provide numerous opportunities for students to be involved in the life of the project. To demonstrate this commitment, the PI has already secured administrative permission from Oxford for the necessary reduction in teaching and service load should the award be made. The budget does not include any teaching buyout.

The PI has an experienced sense of what it will take to manage this research agenda. The PI has a successful management track record with large projects funded by public scientific agencies, such as the multi-year 1.23 million USD, 3 year grant from the US NSF to study "The Effect of the Internet on Society: Incorporating Central Asia into the Global Perspective" (NSF Award ITR-0326101).

Additional Resources

In addition, PI Howard has the management support team of the Oxford Internet Institute. Their assistance coordinating and managing the day-to-day administration of the project will make the COMPROP's research objectives achievable. The administrative staff will assist in legal and contracting matters, ICT support, including administration of the COMPROP website and data systems. These staff positions have already been filled and their responsibilities will be restricted to activity coordination. Also, the OII has the space to support meetings, workshops and regular interaction between the colleagues on the project.