Defending the Democratic Open Society in the Cyber Age – Open Data as Democratic Enabler and Attack Vector

Dr. Jan Kallberg
Captain W. Blake Rhoades
Cadet Marcus J. Masello
Dr. Rosemary A. Burk

In the security paradigm, privacy is the major challenge for the security of an open society against cyber threats. In contemporary society, privacy is a lesser security challenge than the threat of an increased attack surface and strengthened attack vectors: Big Data, artificial intelligence, and the massive aggregation of public data. In this research note, we introduce a high-level conflict between interests and societal goals that supersede the privacy and security conflict.

This conflict is between maintaining an open, democratic society with access and dissemination of digital, public information while concomitantly maintaining security. Dissemination of information can create weaknesses primed for cyberattacks by allowing adversaries access to data. Our intention with this research note is to visualize the problem, assess how it can be addressed, and give a direction for future research.

THE DEMOCRATIC OPPORTUNITY WITH OPEN DATA

As a visualization of the democracy-secrecy dichotomy, we turn to Open Data. The voluntary dissemination of public sector information by the government to include Open Data initiatives are intended to strengthen the democracy, lower costs, and increase a societal understanding of the public sector through transparency and accountability. By releasing massive datasets, the government can be studied in detail. Democratic doctrine assumes that, by default, it is beneficial for the constituency to be well-informed, to have access to primary knowledge of the public sector, and that resources entrusted to the public sector are utilized properly. As a democracy, it is pivotal to seek

Dr. Jan Kallberg is Assistant Professor of American Politics in the Department of Social Sciences and Cyber Policy Fellow at the Army Cyber Institute at West Point. He holds a Ph.D. in Public Affairs and a Master's of Political Science from the University of Texas at Dallas; and a JD/LL.M. from Stockholm University. Prior to joining the West Point faculty, Jan was a researcher and Post-Doc at the Cyber Security Research and Education Institute, Erik Jonsson School of Engineering and Computer Science at the University of Texas at Dallas under Dr. Bhavani Thuraisingham.

Dr. Kallberg's research interest is the intersection between public leadership and cyber capabilities; especially offensive cyber operations as an alternative policy option. His personal website is www.cyberdefense.com.

the consent of the governed, and the governed has, by democratic doctrine, to understand governance and the utilization of public resources. The consent of the people is a foundation for the legitimacy and accepted authority of a democratic republic.

President Lincoln stated in a speech in 1854:

> I have quoted so much at this time merely to show that, according to our ancient faith, the just powers of governments are derived from the consent of the governed. Now the relation of master and slave is pro tanto a total violation of this principle. The master not only governs the slave without his consent, but he governs him by a set of rules altogether different from those which he prescribes for himself. Allow all the governed an equal voice in the government, and that, and that only, is self-government. [1]

President Lincoln's speech was not unique; it followed a philosophical tradition from Aristotle, Locke, Jefferson, and forward, who put forth that citizenry of a republic could only succeed if it was engaged and knowledgeable of how society worked. The United States Declaration of Independence says, "That to secure these rights; governments are instituted among men, deriving their just powers from the consent of the governed." Consent from an uninformed constituency is not actual consent and does not contribute to a democratic process, so it has no value as a vehicle for the legitimacy of the republic. The core concept of the democratic republic is that the people will elect representatives based on merit and trust, for the betterment of the people, and that the elected representatives carry out the people's public business as intended by the governed.

Ignorance and lack of knowledge undermine the legitimacy of the democratic republic. Instead of

Captain Blake Rhoades is a member of the Army Cyber Institute's Innovation Team and an Instructor of International Relations in the Department of Social Sciences. From 2012-2013, he was the company commander of the Army's first Cyber National Mission Team at the 780th MI Brigade in Ft. Meade, MD, and has deployed twice as a signals intelligence platoon leader in support of Operation Iraqi Freedom. Blake holds an M.S. in Information Security Policy and Management from Carnegie Mellon University and a B.A. (Political Science) from the University of Alabama. CPT Rhodes recently served as a Madison Policy Forum cybersecurity fellow in New York, NY.

people being governed by fellow members of the republic, they are governed by a faction supported by procedures and empty mechanics. Early on, the Founding Fathers identified the crucial impact of openness for a functional democratic republic, visualized by Thomas Jefferson in his quote, "An informed citizenry is at the heart of a dynamic democracy."

OPEN GOVERNMENT

If we want a professional government and a functional democracy, we cannot surrender the leadership of the republic to bureaucrats. The desire to create an open government with higher accountability, transparency, and efficiency has grown over time and could be seen as a product of our professionalized federal government where the citizenry is the principal, directly or through their elected officials, and the professional public administration is the agent.[2]

US government initiatives to disseminate digital information accelerated in the 1990s during President Clinton's administration,[3] continued under President Bush, and received strong support in the early President Obama administration.[4][5] The Federal Office of Management and Budget (OMB) Open Government Directive has outlined a set of principles for Open Data dissemination:

> **In general, open data will be consistent with the following principles:** Public. Consistent with OMB's Open Government Directive, agencies must adopt a presumption in favor of openness to the extent permitted by law and subject to privacy, confidentiality, security, or other valid restrictions.
>
> **Accessible.** Open Data are made available in convenient, modifiable, and open formats that can be retrieved, downloaded,

Marcus J. Masello is a senior Army ROTC Cadet studying Information Technology at the University of Toledo. At his battalion, he is the Cadet S6 Communications Officer, and on campus participates in the Association of Information Technology Professionals where he is the Director of Membership for its Toledo chapter. Originally from Youngstown, Ohio, he earned his Eagle Scout Award in 2011 and graduated from Boardman High School in 2014. Last summer, Cadet Masello was selected to intern at the Army Cyber Institute at West Point and upon graduation hopes to branch active duty Cyber Corps.

indexed, and searched. Formats should be machine-readable (i.e., data are reasonably structured to allow automated processing). Open Data structures do not discriminate against any person or group of persons and should be made available to the widest range of users for the widest range of purposes, often by providing the data in multiple formats for consumption. To the extent permitted by law, these formats should be non-proprietary, publicly available, and no restrictions should be placed upon their use.

Described. Open Data are described fully so that consumers of the data have sufficient information to understand their strengths, weaknesses, analytical limitations, security requirements, as well as how to process them. This involves the use of robust, granular metadata (i.e., fields or elements that describe data), thorough documentation of data elements, data dictionaries, and, if applicable, additional descriptions of the purpose of the collection, the population of interest, the characteristics of the sample, and the method of data collection.

Reusable. Open Data are made available under an open license that places no restrictions on their use.

Complete. Open Data is published in primary forms (i.e., as collected at the source), with the finest possible level of granularity that is practicable and permitted by law and other requirements. Derived or aggregated open data should also be published but must reference the primary data.

Timely. Open Data are made available as quickly as necessary to preserve the value

Dr. Rosemary Burk is a Senior Biologist with the U.S. Fish and Wildlife Service, Ecological Services Division in Pacific Northwest Region. She earned a Ph.D. in Biology from the University of North Texas with a specialization in aquatic ecology and environmental science. She has co-authored several articles that have linked failed cyber defense and environmental consequences including *Failed Cyberdefense: The Environmental Consequences of Hostile Acts,* which was published by U.S. Army Military Review in 2014.

> of the data. The frequency of release should account for key audiences and downstream needs.
>
> **Managed Post-Release.** A point of contact must be designated to assist with data use and to respond to complaints about adherence to these open data requirements. [6]

These initiatives have proliferated into state and local government practices including public utilities and other services that are public assets. Further aims of government's online activity are to serve citizens and bring government closer to the people. The Internet empowers people through transparency, e-voting, collecting opinions on public matters, and increasing political self-efficacy among citizens. Since knowledge of the future is unknown, researchers create scenarios for the future state of e-government[7]; the key question is whether the Open Data increase accountability and transparency. The amount of information the government can publish is immense; however, the publication itself does not automatically translate to trust and confidence from citizens. Open Data can also be a proxy for democracy and bring the government closer to the citizenry. According to its proponents, e-government increases efficiency in service offerings and saves money for the public sector. [8]

The four ways of disseminating public information described by Suzanne Piotrowsk[9]–public meeting, leaks, voluntarily dissemination and freedom of information request—are driven by other actors than the bureaucracy itself. Piotrowski sees this information sharing as part of the political processes. The voluntary dissemination, which freely accessible Open Data would be, historically has rarely been seen at a global level until recent years. The voluntary dissemination is a political decision. The first countries and states in a federal framework to actively pursue dissemination enabling citizens' access to Open Data were mainly the US, Canada, UK, Australia, and New Zealand. One reason these countries are more active in dissemination could be the conflict between bureaucratic interest and the interest of the civic societies where Anglo-Saxon countries have a weaker bureaucratic culture in comparison to political structures in centralized governments.[10]

MAINTAINING LEGITIMACY IN A DIGITAL WORLD

Legitimacy concerns not who can lead but who can govern.[11] Dwight Waldo believed that we need faith in government for it to have a strong legitimacy; it has to protect, deliver, and promise that life will be better for its citizens. With his long career as a political scientist, Waldo conducted comparisons over several decades. He noted, "a massive amount of evidence indicates a decline in traditional sources and loci for legitimacy."[12] Waldo raised the question that if the central glue that holds society together is the expectation of more, what does that lead to? Waldo meant that if we build our society around a government that always delivers more services, benefits, and progress, what would happen if there were less of everything in the future? People need a sense that they are represented, and that government is working to improve their lives. In eras of internationalization and globalization, Waldo predicted that government cannot isolate itself from world events.

The idea that internationalization and globalization undermine legitimacy by creating a blurred political landscape is a theme that Robert A. Dahl voices in his book *On Democracy.* Increasing complexity and distance from the population that exists in international organizations, trade agreements, and bilateral agreements play a role in politics and decrease legitimacy; citizens lose the sense that government actions are in the interest of the people. In the "Administrative State," Waldo defined his vision of the "good life" as the best possible condition for the population that can be achieved based on the time, technology, and resources. A legitimate government demonstrates to its citizens that taxes are not collected then squandered and that the return on the taxes makes them worth paying. The government proposes to the population that it can do a better job for all citizens and the charge for those services is taxation. The dissemination of public information becomes instrumental in upholding legitimacy of the government and enables trust in government during difficult times. If the government is no longer considered legitimate, our government and society have failed.

THE ATTACK VECTOR

Open Data releases can appear inconsequential one by one. When taken collectively, the significance of the Open Data can be exploited by adversaries, though the data itself may provide insight into attack vectors. The U.S. Geological Survey (USGS) publishes data regarding water flow, water volume, and measurements from numerous measuring stations throughout a watershed.[13] National Weather Service (NWS) delivers open data weather information.[14] The U.S Army Corps of Engineers (USACE) provides detailed information about dams, critical levels, and flow.[15] The inferences of this data provide insight into attack vectors. In addition to this open information, the USACE's detailed database in the National Inventory of Dams (NID) has historically been hacked and compromised.[16] An adversary can utilize this information to harm the US in a large-scale cyberattack to destabilize the integrity of dams through a watershed.[17][18] This is a single but compelling example, and we have several others that will be a foundation for our future research.

DAMNED IF YOU DO, DAMNED IF YOU DON'T

The US needs transparency to survive as a society and democracy, but how do we do that without creating an unprecedented cyberattack vector into the core of our community? A problem with Open Data is not a single data source by itself, but the aggregated knowledge conceived by mashing data volumes and creating views and understandings beyond the current state.

Then the question arises, how this can be mitigated so the 'open' constitutional democracy can maintain its democratic posture and still avoid the creation of a broad attack vector. In the initial study, there are four potential researchable approaches, each of them with their strengths and weaknesses.

APPROACHES TO SOLVING THE DEMOCRACY-SECRECY DICHOTOMY

CONCEPT	SUCCESS-LIMITING FACTORS	INCREASED INSECURITY
Security review before release of Open Data sources.	Requires that you understand the adversarial intent and ability as well as the adversary—which is unlikely.	The security review cannot be one data source at a time, but instead the effect of utilizing several sources. A roadmap for attacks is created in this process.
Strike an equilibrium by assigning metrics for vulnerability and democratic value and run it through a risk model.	First, it is a normative process. Second, the Constitution is not a grayscale where you can pick a place on the scale. You are either constitutional or not.	This model generates less insecurity because it is at a high-level.
Limit security concerns based on a resilience assessment and the rapid responses to patch vulnerabilities in our market economy. The approach is similar to the armoured warfare concept of protection through mobility instead of hardening.	The assumption is that the free market economy is quick to patch vulnerabilities and that any damage can be rapidly contained and mitigated. This would favor the dissemination as a considerable benefit to society than the actual risk. The risk is that the assumption is untested.	The increased insecurity is the risk that the underlying assumption fails. If the assumption fails, then the approach is a passive stance enabling an adversary added target vectors and options.
Open Data is centralized, and all releases are from one major repository, which enables an ongoing risk assessment and ability to limit release if necessary.	Once data is released to the public domain, it cannot be recalled.	The risk is a one-stop-shop for data that the adversary can leverage.

We are early in the learning curve and have not thoroughly researched or addressed the security concerns of Open Data. Initially, the Democracy-Secrecy Dichotomy as it relates to Open Data dissemination needs to be a primary inquiry. How do we strike a balance between living in an Open Society and protecting citizens from the harmful release of data? What can we do to meet both goals? Is there a systematic approach that can be applied? The second wave of inquiry is tailored to address case studies and increase the granularity of the research.

The views expressed herein are those of the authors and do not reflect the official policy or position of the Army Cyber Institute, the United States Military Academy, the Department of the Army, the Department of Defense, the United States Fish and Wildlife Service, the Department of the Interior or the United States Government or the University of Toledo.

NOTES

1. Abraham Lincoln, speech at Peoria, Illinois, in Reply to Senator Douglas (October 16, 1854), published in *The Complete Works of Abraham Lincoln* (1894) Vol. 2.

2. Wallace Park, "Open Government Principle: Applying the right to know under the Constitution." Geo. Wash. L. Rev. 26 (1957), 1.

3. Patrice McDermott. "Building Open Government," *Government Information Quarterly* 27, no. 4 (2010), 401-413.

4. The White House, "Transparency and Open Government," Memorandum for the Heads of Executive Departments and Agencies (2009), https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2009/m09-12.pdf.

5. Dennis Linders and Susan Copeland Wilson, "What is Open Government? One Year after the Directive," in *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times,* ACM, 2011, 262-271.

6. U.S. Government, Project Open Data, Open Data Policy–Managing Information as an Asset, https://project-open-data.cio.gov/principles/.

7. Katleen Janssen, "The influence of the PSI directive on open government data: An overview of recent developments," *Government Information Quarterly* 28, no. 4 (2011), 446-456.

8. Jan Kallberg, "The Internet as a Proxy for Democratic Accountability and Transparency---a Comparative Test of Waldo's Five Problem Areas in Five Advanced Democratic Societies," Dissertation, (Richardson: The University of Texas at Dallas, 2011).

9. Suzanne J. Piotrowski, *"Governmental Transparency in the Path of Administrative Reform"*, (New York: SUNY Press, 2007).

10. Paul G. Mahoney, "The Common Law and Economic Growth: Hayek might be Right," *The Journal of Legal Studies* 30, no. 2 (2001), 503-525.

11. David H. Rosenbloom, *"Revisiting Waldo's Administrative State: Constancy and Change in Public Administration"*, (Washington DC: Georgetown University Press, 2006).

12. Dwight Waldo, *The Enterprise of Public Administration: A Summary View,* (Novato, CA: Chandler & Sharp Publishers, 1980).

13. U.S. Geological Survey, Water Watch, https://waterwatch.usgs.gov/.

14. National Weather Service http://www.weather.gov/.

15. U.S Army Corps of Engineers (USACE) http://water.usace.army.mil/.

16. Bill Gertz, "FBI Eyes Chinese Hacking of Dams Database", *Washington Times,* January 6, 2015, http://www.washington-times.com/news/2015/jan/6/fbi-eyes-chinese-hacking-of-dams-database/.

17. Rosemary A. Burk, and Jan Kallberg, "Cyber Defense as a part of Hazard Mitigation: Comparing High Hazard Potential Dam Safety Programs in the United States and Sweden," *Journal of Homeland Security and Emergency Management* 13, no. 1 (2016), 77-94.

18. Jan Kallberg, and Rosemary A. Burk, "Failed Cyberdefense: The Environmental Consequences of Hostile Acts," *Military Review* 94, no. 3 (2014), 22.