Data for Public Policy in the Web Era: A Call for Systemic and Ethical View

Aleksey V. Turobov
PhD student, Doctoral School of Political Science
National Research University Higher School of Economics
Moscow, Russia

Data in and for Public Policy: Why It Is More Important Than It Seems

The growing technical and analytical capabilities of working with data and the active development and expansion of the data science industry obscure a problem arising in governments and law enforcement, which is to be taken seriously if the industry is to persist and prosper. This is problem of how to ensure the effectively use of data and provision of employees with useful information, rather than to encourage a simple accumulation of data which only strengthens the bureaucracies and rising costs. Untimely received data, or data coming in a too large volume without structuring and systematization, as well as meagre, non-informative data, can become an obstacle for governments agencies and may ultimately result in total neglect of the "signs on the wall". Provided that modern technical methods and reasoned methodology correctly collect, systematize, and analyse and provide data to policy makers and bureaucrats, the information obtained can shed light on unsolved crimes, high-risk sites and police actions. Obviously, technological progress has reduced the costs of data collection and storage, and these data have become more accessible to the public, policy makers and the law enforcement services.

In recent decades, there has been a growing understanding that data "can reduce uncertainty about the best course of action"[1] in policy development, that is, it can help to produce a more effective policy-making process and lead to more adequate solutions and more effective public policies. Therefore, political scientists,

---

[1] Data visualisation: Contributions to evidence-based decision-making. A SciDev.Net Learning Report URL: (https://social.shorthand.com/SciDevNet/3geA2Kw4B5c/data-visualisation-contributions-to-evidence-based-decision-making)

policy makers and policy advocates are often inclined to provide evidence-based arguments for specific policy decisions usually obtained through reasonable empirical research or analysis on this topic. The aforementioned technological trends will largely determine the future of public administration and, ultimately, politics. Governments are becoming increasingly aware and adaptable to these trends, and the data science is beginning to play an important role in public policy. In addition, there is a growing understanding that these trends should be addressed by civil society in order to ensure a more democratic governance, accountability and influence on policy development by non-institutional actors. Accordingly, initiatives such as the Civic Analytics Network[2] and the Data Science for Social Good[3] contribute to bringing the data science to solve real-life problems of communities, to provide the necessary tools or technical infrastructure, including training professionals, to empower the civil society in policy development. In other words, these initiatives are aimed at helping societies to overcome technological and information barriers in order to participate in policy development in the modern data environment.

A new "landscape" of data opens up opportunities not only for governments, but for other stakeholders involved in political processes. While governments are already taking steps to integrate data science into governance, non-state actors will have to take this trend more seriously in the near future and adapt their functions and capabilities so that they can respond to policy issues through analysis and integrated data sets. In this regard, every political actor will need to pay more attention to building up its data science capacities. This ultimately means that data science will play an increasingly important role, along with data analysts, in informing the sphere of public policy and formulating policy proposals conforming with the standards of evidence based policy. These goals will require more advances in the science of data. However, the statistical-mathematical algorithms, machine learning, neural networks, etc. as a foundation for political decision-making and the framework of governance

---

[2] Harvard Initiative URL: https://datasmart.ash.harvard.edu/civic-analytics-network

[3] Chicago University Initiative URL: https://dssg.uchicago.edu

and political management are no longer the manifestations of wishful thinking. Instead, they are what can become real if we are serious to let them in the political realm.

Almost any discussion on data in policy making and politics starts within a specific frame that has to do with the social media and the mass media. While consumers often drive social media content, policymakers and governments can also become powerful forces on platforms like Facebook, Twitter, and Instagram. For instance, the National Science Foundation (NSF) created a social media policy[4] for just this purpose. According to the NSF, the organization uses social media to "engage with our stakeholders and with the public." While public organizations might use social media to gauge the public's interest in a specific policy, citizens have become more likely to educate themselves on political issues through the social media.

The Pew Research Center[5] conducted recently a survey on politics and the social media. Thirty-six percent of respondents who use social media said that social media remained between "somewhat" and "very important" when it came to consuming news about political issues. As citizens turn to social media to learn about politicians and to debate political issues, policymakers need to pay more attention to the ways in which information is disseminated to the public.

Additionally, politicians can find new ways to engage with the public on policy issues affecting citizens. Over the past few years, politicians have become more active on the social media. Reporter Philip Bump wrote[6] in the Washington Post, for example, that, since he announced his candidacy, Donald Trump had never spent more than 46 hours without at least one tweet. When politicians use the social media, they can influence public policy or engage with their constituents to better understand

4 National Science Foundation. Social Media: Comment Policy, Disclaimer, Privacy, Copyright URL: https://www.nsf.gov/social/policies.jsp

5 Pew Research Center. Social and Demographic Trends. Rew Charitable Trust URL: http://www.pewsocialtrends.org/pew-charitable-trusts/

6 The Washington Post. The only information Trump supports is information that makes him look good URL: https://www.washingtonpost.com/news/politics/wp/2018/04/03/the-only-information-trump-supports-is-information-that-makes-him-look-good/?utm_term=.60617de9fd58

them. A recent eMarketer survey has reported[7] that nearly 80 percent of adult internet users feel that the social media have at least some influence over public policy decisions. The website cited immigration and free trade as two examples, but other forms of public policy can also come under scrutiny on the social media. According to eMarketer, however, some Americans believe that the social media enjoy a positive impact on politics and public policy, while others feel more skeptical. Leaders of today and tomorrow might need to lead citizens toward positive social media usage when it comes to discussing public policy.

This phenomenon is crucial for modern political science's understanding of the ongoing transformational processes in politics. Interestingly, digital space has so far recorded more vibrant criticism in the social media, chat rooms, or messenger programs in both liberal and illiberal regimes, although the exact expression of this dissent yields different explanations in numerous studies [Perea, Jensen, Jorba 2013]. Pippa Norris [2002, 2012], for example, demonstrated how participation in e-petitions, protests, and sit-ins had been altered through digital communication technologies, although her later work with Ronald Inglehart demonstrated how political polarization in the Western democracies largely resulted in a backlash against the much prophesized digital democracy argument. In a study on American, Australian, and British young voters, political use of social media directly correlates with offline political activism, with similar patterns of polarization observed both online and offline (media use and youth political engagement in three advanced democracies) [Xenos, Vromen, Loader 2014]. In another study, citizen-initiated campaigning outside the United States is observed as a political participation method largely benefiting the hegemonic party with little advantage to opposition movements [Gibson 2015]. The relationship between political engagement and social media use is still being challenged, however, as numerous studies have found mixed results

---

[7] eMarket survey report
https://www.emarketer.com/Article/Most-Internet-Users-Prefer-Search-Engines-Find-Local-Products/1015737

related to how online and offline types of political participation work in tandem [Zúñiga, Molyneux, Zheng 2014, Bode, Vraga, Borah, Shah 2014].

The rapid development of computing power and digital technologies is therefore coming to the attention of states, political institutions and actors [Cornish 2010] and is, in turn, becoming an independent object of research interest in political science [Tufekci 2014; Wanna 2018; Hecker, Haklay et al. 2018].

The fact that political science today is short of a valid and scientifically sound approach to data, new era of information, the social media and the mass media on internet and so on, is a real quality driver for (re)creating an interdisciplinary space to study social and political processes under conditions of technological advances. This is extremely important, because all these issues are directly related to the construction of e-government and digital transformation of public governance, especially during the period when states publicly announce the development or adoption of national digitalisation programs.

Discussion and conclusion

Now we come to the debatable questions of errors of the existing models of data usage, including algorithms and the methods, as well as techniques of their improvement. One of the most serious methodological errors is the 'use of specific data' - a set of data stored with systematic bias (models will reproduce and in some cases reinforce the same biases). In a best case scenario, this leads to the fact that prognostic models are ineffective. In the worst case, this leads to results getting completely discredited. In this way, when we are planning a scientific study, we should always think about the reliability and data engagement. Certainly, we need to understand ethical and moral issues when we proceed to the interpretation of results based on data of this kind. Similarly, when discussing questions about data, it should be noted that it is advisable to expand the sources of data acquisition [Saunders, J., Hunt, P., & Hollywood, J. S. 2016].

Another problem of constructing any model in discourse of public policy is the collection and extraction of qualitative data. It is established that all the data used are

distributed across different databases, are difficult to structure, or are not practically systematised. It is necessary to collect and combine data from various statistical agencies and government institutions to be able to use current data for future analysis and making evidence based decision in politics, etc. It is necessary to automate systems that can combine data from various resources, quickly classify and make available for use. At the same time, governments, public and social institutions and law enforcement bodies need better information about what to do with the result of such automation analysis as part of decision-making process.

These two problems are in fact a part of a fundamental question: what kind of relationship is there between the Digital Tech and politics? Is technology building rules (or providing imperative suggestions) for governments? Or is it politics that drives technology growth and controls it? Is there a possibility for the Digital Tech and Politics to build a new relationship on a parity basis? We do not have a definite answer. But we can (or rather should) start thinking about ethical and moral issue of applying digitalisation to produce more knowledge in order to build a more ethical relationship.

How can we start working on it? Let us think about shaping a systemic view on potential perspectives:

"Positive" hypotheses of impact in relationship "Policy <-> Digital technology"	"Negative" hypotheses of impact in relationship "Policy <-> Digital technology"

In the presence of a well-grounded theoretical and methodological justification and a valid empirical study, the formation of the evidence based policy, using data analysis and mathematical-statistical algorithms (next step data usage) directly in the sphere of political decision-making, can "exclude the human factor" (emotions, personal interest, corruption, etc.) from politics, which will prevent and control the emergence and development of authoritarian policies, concepts and ideas	The lack of decision-making practice on the basis of using automated systems (mathematical-statistical algorithms) raises the problem of changing the subject of political decision-making (when the political actor makes decisions constantly, how the subjective perception of the fact of decision making by the algorithm will change and how many actors will trust such decisions)
The achievements of the scientific and technical process (increasing technological capacities, creating new ways of analyzing data, etc.) allow us to process large amounts of data and, as a consequence, to construct objective solutions and formulate an objective statement of goals and objectives	All existing mathematical-statistical models and algorithms build forecasts (predict) solely on the basis of previous experience. In other words, if a situation emerges not from the developed "norm" (the norm in this case is formed on the basis of the previous experience - data) - the automated system will give an erroneous answer

Automated systems for political decision-making remove the psycho-emotional and personal factors of the burden of responsibility of a political actor / organization	Analytically and practically controversial, the question of the distribution of responsibility arises: who is responsible if an automated system has made a decision causing a considerable damage? (how to allocate responsibility between the ideologists of an automated system, system engineers, political actors implementing this system, etc.)
Algorithms and technological models make it possible to "depersonalize" politics (the exclusion of cults of personalities, etc.)	The "depersonalization" of politics can lead to extremely negative consequences expressed in the loss of dialogue with the public (the emergence of the view that a person is needed only for "trivial" tasks or for creative tasks far from political governance)

* * *

Though digital technologies are yet to become fully mature, it is clear that they are here to stay. They can help, but they are also a source of numerous problems. Letting computers in the public policy domain demands thinking in, alas, mutually contradictory terms of "cost effectiveness" and "rights and freedoms". The arguments of expediency rising from the digitalization and automation capacity to generate immense profits and provide security run into moral and ethical arguments.

Moral and ethical issues arise on almost every stage of data usage. "Bad data" result in outcomes leading to discrimination; inconclusive evidence leads to unjustified actions; inscrutable evidence leads to opacity; misguided evidence leads to bias [Mittelstadt, Allo, Taddeo, Wachter, Floridi 2016].

Almost any algorithm has an impact. So, companies producing algorithms should be accountable for what they do and how they do. It means that the development of ethical codes and moral rules are to become an integral part of the data and algorithms industry. To achieve this goal we need a "communication bridge" between researchers (academics), companies and governments. Really smart algorithms are likely to result from interdisciplinary studies. Algorithms doing good in a fashion prescribed by utilitarianism are likely to result from the sensitivity to moral and ethical issues of all the parties involved. However, establishing what is good is a tricky business. If we are not to drown in the interplay of private and group preferences, the only viable option has to do with an evidence based approach and transparent public policy. The former is a tool, while the latter is the sphere where socially acceptable goals can be worked out with an eye on moral and ethical issues. Technology can produce a lot of good or enormous harm. Probably, it is the right time to start learning by our previous mistakes in order not to become empowered by algorithms and data instead of getting enslaved by them.

References:

1. Eva Anduiza Perea, Michael James Jensen, and Laia Jorba, eds., *Digital Media and Political Engagement Worldwide: A Comparative Study*; Jennifer Oser, Marc Hooghe, and Sofie Marien, "Is Online Participation Distinct from Offline Participation? A Latent Class Analysis of Participation Types and Their Stratification," *Political Research Quarterly* 66, no. 1 (2013), 91-101.

2. Pippa Norris, *Democratic Phoenix: Reinventing Political Activism* (Cambridge: Cambridge University Press, 2002); Ronald Inglehart and Pippa Norris, "Trump, Brexit, and the Rise of Populism: Economic Have-Nots and Cultural Backlash" (Harvard Kennedy School Working Paper No. RWP16-026, Harvard University, Boston, MA: 2017) (available at SSRN: https://ssrn.com/abstract=2818659)

3. Michael Xenos, Ariadne Vromen, and Brian D. Loader, "The great equalizer? *Patterns of social & Society* 17, no. 2 (2014), 151-167.

4. Rachel K. Gibson, "Party change, social media and the rise of 'citizec-initiated' campaigning" Party Politics 21, no. 2 (2015), 183-197

5. Homero Gil de Zúñiga, Logan Molyneux, and Pei Zheng, "Social Media, Political Expression, and Political Participation: Panel Analysis of Lagged and Concurrent Relationships," *Journal of Communication* 64, no. 4 (2014), 612-634

6. Leticia Bode, Emily K. Vraga, Porismita Borah, and Dhavan V. Shah, "A New Space for Political Behavior: Political Social Networking and Its Democratic Consequences," *Journal of Computer Mediated Communication* 19, no. 3 (2014), 414-429;

7. Nakwon Jung, Yonghwan Kim, and Homero Gil de Zúniga, "The Mediating Role of Knowledge and Efficacy in the Effects of Communication on Political Participation," *Mass Communication and Society* 14, no. 4 (2011), 407-430.

8. Paul Cornish (2010) Technology, strategy and counterterrorism. International Affairs (Royal Institute of International Affairs 1944-), Vol. 86, No. 4 (July 2010), pp. 875-888

9. Zeynep Tufekci (2014) Social Movements and Governments in the Digital Age: Evaluating a Complex Landscape. *Journal of International Affairs*, Vol. 68, No. 1, Breaking Point: Protests and Revolutions in the 21st Century (FALL/WINTER 2014), pp. 1-18

10. John Wanna (2018) Opening government: Transparency and engagement in the information age. Book Title: Opening Government. Book Subtitle: Transparency and Engagement in the Information Age Book Editor(s): JOHN WANNA, SAM VINCENT. Published by: ANU Press

11. Susanne Hecker, Muki Haklay, Anne Bowser, Zen Makuch, Johannes Vogel and Aletta Bonn (2018) Innovation in open science, society and policy – setting the agenda for citizen science. Book Title: Citizen Science. Book Subtitle: Innovation in Open Science, Society and Policy. Book Editor(s): Susanne Hecker, Muki Haklay, Anne Bowser, Zen Makuch, Johannes Vogel, Aletta Bonn. Published by: UCL Press. (2018)

12. Saunders, J., Hunt, P., & Hollywood, J. S. (2016). Predictions put into practice: a quasi-experimental evaluation of Chicago's predictive policing pilot. *Journal of Experimental Criminology*, *12*(3), 347–371.

13. Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter and Luciano Floridi (2016) The ethics of algorithms: Mapping the debate. Big Data & Society. July–December 2016, 1–21.