Promises and Pitfalls of Artificial Intelligence for Democratic Participation

Workshop Proceedings

CCDSEE, GSI, University of Geneva, December 10 – 11, 2020, Virtual Event

Abstract

When discussing the role of digital technologies in pluralist democracies, artificial intelligence (AI) is probably the most controversial. This technology triggers the greatest hopes (e.g. AI can help making sense out of millions of comments) and at the same time raises many concerns (e.g. AI can reinforce existing biases). As AI remains quite difficult to grasp, including for those in a position to make a political decision (i.e., policy makers and citizens), the main objective of this workshop is to unpack the current promises and challenges associated with AI for democratic participation. This report presents a summary of the authors' contributions. The first session discusses conceptions of AI (the gap between reality and expectations, education challenges, and media frames). Session 2 considers promises and pitfalls of AI to foster citizen mobilization. Session 3 explores how AI can support collective intelligence processes, including civic tech. Session 4 discusses how AI transforms the role and the making of citizens. Session 5 presents key promises of AI for governments.

Keywords: Artificial intelligence, AI, Digital technologies, Political participation, Citizen participation, Democracy, Europe.

Organizers

- Dr. Jérôme Duberry, Research Fellow, Dusan Sidjanski Centre of Excellence for European Studies, Global Studies Institute, University of Geneva.
- Sabrya Hamidi, Research Assistant, Dusan Sidjanski Centre of Excellence for European Studies, Global Studies Institute, University of Geneva.

Corresponding author: Jerome.duberry@unige.ch

Introduction

> "In late January, researchers at BenevolentAI, an artificial intelligence start-up in central London, turned their attention to the coronavirus. Within two days, using technologies that can scour scientific literature related to the virus, they pinpointed a possible treatment with speed that surprised both the company that makes the drug and many doctors who had spent years exploring its effect on other viruses."[1]

This excerpt from The New York Times discusses the role of artificial intelligence (AI) in a society deeply affected by the emergence of the COVID-19. The quote illustrates how AI can support the work of scientists to find a cure and halt the spread of the pandemic. AI is indeed increasingly used in our everyday life, including in politics, art, and security to name only few. This technology performs a wide range of activities, including "verbal-linguistic, visual-spatial, logical-mathematical, naturalistic, and interpersonal intelligence"[2]. Since AI is still under development, and its uses are extremely diverse, its conceptualization and definition remain a challenge. As Monett and Lewis argue, "[t]heories of intelligence and the goal of Artificial Intelligence (A.I.) have been the source of much confusion both within the field and among the

[1] Cade Metz. 2020. How A.I. Steered Doctors Toward a Possible Coronavirus Treatment. The New York Times. Retrieved from https://www.nytimes.com/2020/04/30/technology/coronavirus-treatment-benevolentai-baricitinib.html

[2] Dagman Monett et al, (2020). Special Issue "On Defining Artificial Intelligence" —Commentaries and Author's Response" Journal of Artificial General Intelligence. Volume 11, Issue 2, p.19

general public"[3]. Hence, politicians, regulators and civil society must acquire a better understanding of this technology[4] and the associated hopes and concerns it triggers. The risk is to approach AI with fear and suspicion, which may lead to "slow AI's development or drive it underground (…), impeding important work on ensuring the safety and reliability of AI technologies"[5].

This workshop focuses on the use of AI in the context of democratic participation. Each contribution sheds a different light on the promises and concerns raised by the increasing use of AI in this specific context. We decided to adopt an operational definition of AI, which highlights the combination of two criteria that distinguish between AI and computer systems: information processing and adaptation. Wang defines AI as "the capacity of an information-processing system to adapt to its environment while operating with insufficient knowledge and resources"[6]. He argues that the information processing capacity of AI consists of choosing and executing tasks, and adjusting its behavior according to its past experiences[7]. This definition fits particularly well our discussion about the promises and pitfalls of AI for democratic participation since it enables us to better distinguish between AI and former tech-related issues.

Goal(s) of the workshop

The concept of Augmented Democracy (AD), developed by Hidalgo and his team[8] at the Massachusetts Institute of Technology (MIT) proposes to explore how AI could foster democracy. It is based on the assumptions that (1) more information and more instances of participation are not necessarily the solution to the decline of citizen participation in pluralist democratic states, and that (2) AI is progressively infiltrating all aspects of our lives, including politics. When discussing the role of digital technologies in pluralist democracies, AI is probably the most controversial. This technology triggers the greatest hopes (e.g. AI can help making sense out of millions of comments) and at the same time raises many concerns (e.g. AI can reinforce existing biases). This workshop focuses on the use of Artificial Narrow Intelligence (ANI). In other words, it focuses on current uses of AI. It does not explore or discuss the future developments of AI, such as Artificial General Intelligence (AGI) or Artificial Super Intelligence (ASI) and the point of singularity. Since AI remains quite difficult to grasp for non-experts, including for those who are in the position of making a political decision (i.e., policy-makers and citizens), this workshop's primary goal is to unpack ongoing promises and challenges associated with AI for democratic participation. It also contributes to bridge the gap between social and computer scientists and policy experts by exchanging knowledge and experiences.

Acknowledgments

We would like to express our warmest appreciation to all the participants of this workshop for their valuable contributions, and their enthusiasm despite the challenging health and logistical conditions. We also would like to thank the Dusan Sidjanksi Center of Excellence for European Studies[9], the Global Studies Institute [10]of the University of Geneva and the Swiss National Science Foundation[11] for their generous support.

[3] Monett, D., and Lewis, C. W. P. 2018. Getting clarity by defining Artificial Intelligence - A Survey. In Muller, V. C., ed., *Philosophy and Theory of Artificial Intelligence 2017*. Berlin: Springer. 212–214.
[4] Al-Amoudi, I., & Latsis, J. (2019). Anormative black boxes: Artificial intelligence and health policy. In *Post-Human Institutions and Organizations* (pp. 119-142). Routledge.
[5] Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., ... & Teller, A. (2016). Artificial intelligence and life in 2030: The one hundred year study on artificial intelligence, p.298.
[6] Wang, P. (2019). On defining artificial intelligence. *Journal of Artificial General Intelligence*, *10*(2), 1-37, p.17
[7] Wang, P. (2019). Op. Cit.
[8] See https://www.peopledemocracy.com
[9] See https://www.unige.ch/gsi/fr/presentation/centres-de-recherche-affilies/le-centre/
[10] See https://www.unige.ch/gsi/en/bienvenue/le-global-studies-institute-gsi/
[11] See http://www.snf.ch/en/Pages/default.aspxa

Agenda

This report presents a summary of authors' contributions. They are organized in five sections. The first session discusses conceptions of AI (the gap between reality and expectations, education challenges, and media frames). Session 2 considers promises of pitfalls of AI to foster citizen mobilization. Session 3 explores how AI can support collective intelligence processes. Session 4 discusses how AI transforms The Political, including the role and the making of citizen. Session 5 presents key promises of AI for governments.

1. Conceptions of AI

Prof. Himanshu Verma: Sociological conception of AI: Bridging the gap between Reality and Expectations
Dr. Alexandre Dormeier Freire: AI in education: fears, gaps and opportunities
Dr. Jérôme Duberry & Sabrya Hamidi: AI definition challenge in the media

Sociological conception of AI: Bridging the gap between Reality and Expectations.

Prof. Himanshu Verma, Delft University of Technology

The way we design AI could be much more grounded in social values, more attuned to societal needs and expectations. The design and deployment of AI is unilaterally controlled and dictated by a selected few in the Silicon Valley -- deciding how and what the AI should look like. Another way of rethinking the design and development of AI is to include multiple stakeholders, and to understand how these stakeholders envision AI to be. One example of such a participatory and bottom-up effort is the co-design of visual signals for driverless shuttles which are running in the city of Sion (Switzerland). Here, the local residents collaborated with Human-Computer Interaction researchers and the service provider of the driverless shuttle to design effective ways of facilitating Pedestrian-AI interaction.

Furthermore, currently there are over expectations in terms of AI's capability and what it can reasonably offer. We know very well that AI is a very effective tool when it comes to solving small problems with a very narrow focus. As soon as we try to scale this up, AI fails miserably, as it is still not perfect. For example, Amazon's face recognition system falsely classified 28 American lawmakers as criminal mugshots.

There is a big gap in the reality of AI and our expectations with AI, because 1) the level of education is low amongst different stakeholders such as Policy Makers, Educators, etc., 2) AI developers are not very knowledgeable about other domains (such as Oncology), and 3) role of other stakeholders in the conception and development of AI is still very low or minimal. So, to bridge this gap, and to have a widely acceptable and sociological conception of AI, we need to work in multi-stakeholder settings where we can ground each stakeholders' needs and expectations with AI, and develop AI with a bottom-up approach. This also entails educating stakeholders about how AI works and what can it reasonably achieve. This was a core idea behind an article that we wrote in 2018 with my colleagues Hamed Alavi, Jakub Mlynar, and Lorenzo Cantoni[12]. The idea was that sociologists so far have seen AI as some kind of a non-human entity which is not to be studied on equal terms with humans. We made a proposition that sociology must play a central role in imagining how AI should be. Also, one of the key messages that came out of the AI Governance Forum (which was co-organized with Dr. Eva Thelisson of the AI Transparency Institute[13], and is also her brain child) is that AI is a force for good but at the same time can be instrument for power. AI ethics should be part of computer science and professional curriculums such as the Hippocratic oath that doctors take.

In our research on the role and impact of AI in clinical research, the way AI is envisioned (expectations and reality do not match), and the way it is used in clinical research is not the same as how it is adopted in clinical practice. Experts also reported that AI will not replace humans and we have to accept this fact. So far, AI has revolutionized oncology in many ways, better sensing techniques, better imaging techniques, better analysis methods, and the capacity to integrate a lot of data, but it's just the tip of the iceberg.

[12] Mlynář, J., Alavi, H. S., Verma, H., & Cantoni, L. (2018, August). Towards a sociological conception of artificial intelligence. In *International Conference on Artificial General Intelligence* (pp. 130-139). Springer, Cham.
[13] https://aitransparencyinstitute.com

One of the reasons why AI has not yet been trustworthily adopted in medical domain is because of this black box approach, and its rate of advancement is too high for it to be tested, validated and adopted (more time for ethics and reliability). But an interesting area to look and this is where AI can help oncologists is with augmenting their cognitive limits. Few steps to improving trust are education, not necessarily about the maths behind AI but we do have to understand what AI does.

AI in education: fears, gaps and opportunities

Dr. Alexandre Dormeier Freire, Graduate Institute of International and Development Studies, Geneva.

The debates of linking artificial intelligence and education goes back to the 80's, if not earlier. There are a certain number of opportunities such as automated tutoring for instance that can free educators from time consuming and repetitive tasks. Teachers would focus more on their core task which is to educate pupils. To summarize, 'smart contents' are basically a way to adapt learning materials. Instead of having a textbook that is quite static, AI can support the individual's path of learning, in a more personalized way, tailored to the needs of individuals, available everywhere in the world, and allowing interaction with the learner. Predictive learning tools can also improve our ability to learn the learning processes themselves.

The purpose to use AI in education is not to replace the teacher but to support its role, it is *blended learning*. Learning is multidimensional, it is not only about cognition but also emotions so obviously when it comes to AI it is questionable, notably about ethics, social biases because AI might just contribute to reproduce social biases within educational systems. Personal data such as our learning curves would be an incredible source of information but also problematic for data security. In addition, we cannot ignore the existence of the digital divide worldwide, accessibility and affordability questions, 80% of the teachers in developing countries are not IT literate. Most of AI models are western, so it raises the question of power relations, on top of the clash between UNESCO vision of a human centric education versus a human-machine learning relation. Curriculums in education are the result of social values, as education reflects the debates of society itself, integrating AI in education is then a problem of defining a position of AI in society. Finally, the level of preparedness of schools and teachers is crucial. The risk of private companies lobbying for the use of certain technologies that do not fit with the mission of education is important. The void left by public actors because they have no financial resources can be quickly filled by the commercial products of private technology companies with possible moral and ethical questions repercussions.

Contrasted media frames of AI during the Covid-19 Pandemic

Dr. Jérôme Duberry, Sabrya Hamidi, CCDSEE, Global studies Institute, University of Geneva.

Traditionally, policy makers and citizens consider mainstream media as a credible source of information. Through agenda setting and framing processes, mass media coverage influences which issues the public are aware of, and what their attitudes towards those issues are.[14] Kitzinger[15] has shown that public attention is correlated to the way media focus on that issue. In the context of pandemic and uncertainty, media plays a key role in "shaping public opinion around emerging science and risk issues, and the degree of politicization and polarization of such news coverage may be important and influential factors"[16]. Previous research explored how news media cover pandemics[17]. Mainstream media remain a trusted

[14] Patterson, C., Semple, S., Wood, K., Duffy, S. and Hilton, S., 2015. A quantitative content analysis of UK newsprint coverage of proposed legislation to prohibit smoking in private vehicles carrying children. *BMC Public Health*, *15*(1), pp.1-7.
[15] Kitzinger J. Framing abuse: Media influence and public understanding of sexual violence against children. London: Pluto Press; 2004.
[16] Hart, P.S., Chinn, S. and Soroka, S., (2020). <? covid19?> Politicization and Polarization in COVID-19 News Coverage. *Science Communication*, *42*(5), pp.679-697, p.680.
[17] See for instance Klemm, C., Das, E., & Hartmann, T. (2016). Swine flu and hype: a systematic review of media dramatization of the H1N1 influenza pandemic. Journal of Risk Research, 19(1), 1-20.

source of covid-19 information[18], as indicated by the sharp rise in the number of online subscriptions to newspapers since the Covid-19: "In the first three months of the year, The New York Times Company added more digital subscribers than it had gained during any quarter since it started charging readers for online content in 2011. But that increase was driven by widespread interest in news of the coronavirus pandemic, which has ravaged the U.S. economy and cut deeply into The Times's advertising revenue"[19].

Our study[20] examined how mainstream media depict AI during a pandemic. We explored how French and English-speaking mainstream media portray AI during a pandemic. We selected five media outlets and extracted all news articles that mentioned AI over a period of 30 days. Our study shows that no news article provided a definition of AI or clarified the relation between the AI related terms such as machine learning or neural networks. Without a clear definition of artificial intelligence, and no indication about the relation between the umbrella term AI/artificial intelligence, and some specific capabilities of this technology, such as machine learning, algorithms, deep learning, neural networks, "it is difficult for policy makers to assess what AI systems will be able to do in the near future, and how the field may get there. There is no common framework to determine which kinds of AI systems are even desirable"[21]. However, we argue that most news items focus on two characteristics of AI that correspond to the operational definition of Wang[22]: processing information and adaptability. In terms of narratives, our research confirms previous studies and shows that the coverage of AI in US mainstream media is "consistently more optimistic than pessimistic."[23] Since many AI application to combat Covid-19 use sensitive data, "more in-depth and concrete discussion of the risks and benefits of AI in news media is needed to allow critical assessment of the use and misuse of AI"[24]. European newspapers of this sample offer a more balanced perspective on AI. They highlight both the benefits and risks associated with AI. The main positive narrative identified is around AI supporting the work of professionals, or said differently, an extension of human capacity. The main negative narrative among the articles of this sample is about ethical issues, and mainly about privacy and surveillance.

Previous studies showed that "[t]he framing of the pandemic changed during the three stages: at first the pandemic was defined as a very high-level risk, then this was brought down to the levels of the common flu, followed again by a renewed concern during the actual epidemic"[25]. This three stage-model offers a possible explanation of the differences in AI media coverage between US and European newspapers. The Covid-19 pandemic did not indeed spread at the same speed in Europe and North America, and political leaders on both sides of the continent did not perceive the seriousness of threat simultaneously. US newspaper frames AI optimistically and presents a mix of applications that are related to art, entertainment, education and health. It does not mainly frame AI according to its benefits and risks to combat COVID-19. In April, Europe experienced the peak of the first wave of COVID-19 pandemic. Consequently, European newspapers depicted AI in the context of a global pandemic. They framed AI's benefits and risks in the context of COVID-19, including ethical issues, data privacy and state surveillance. Hence, this difference of media coverage between US and European media is aligned with past research about media coverage of pandemic.

---

[18] Fletcher, R., Kalogeropoulos, A. and Nielsen, R.K., (2020). News media broadly trusted as source of coronavirus information, views of UK government response highly polarised. *Reuters Institute*.

[19] Tracy, M. (2020). The New York Times Tops 6 Million Subscribers as Ad Revenue Plummets. The New York Times. Retrieved from https://www.nytimes.com/2020/05/06/business/media/new-york-times-earnings-subscriptions-coronavirus.html

[20] Duberry, J., & Hamidi, S. (2021). Contrasted media frames of AI during the COVID-19 pandemic: a content analysis of US and European newspapers. *Online Information Review*.

[21] Bhatnagar, S et al. 2018. Mapping Intelligence: Requirements and Possibilities. In Muller, V. C., ed., *Philosophy and Theory of Artificial Intelligence 2017*. Berlin: Springer. 117– 135.

[22] Wang, P. (2019). On defining artificial intelligence. *Journal of Artificial General Intelligence*, *10*(2), 1-37.

[23] Fast, E., and Horvitz, E. (2016). Long-term trends in the public perception of artificial intelligence. *arXiv preprint arXiv:1609.04904*, p.968

[24] Chuan, C. H., Tsai, W-H. S., Cho, S. Y. (2019). 'Framing Artificial Intelligence in American Newspapers.' *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (AIES '19). ACM, New York, NY, 339–344, p.343

[25] Klemm, C., Das, E., & Hartmann, T. (2016). Swine flu and hype: a systematic review of media dramatization of the H1N1 influenza pandemic. *Journal of Risk Research*, *19*(1), 1-20.

2. AI & Citizen mobilization

Dr. Yanina Welp: Old wine in new bottles? The use of ICTs to promote citizen's engagement
Dr. Jasmine Lorenzini: Supporting youth participation in decision-making processes
Dr. Moritz Büchi: AI and society: a digital inequality perspective
Prof. Diego Garzia: How is AI used in Voting Advice Applications?

Old wine in new bottles? The use of ICTs to promote citizen's engagement in policy-making

Dr. Yanina Welp, Graduate Institute of International and Development Studies

Information and communication technologies (ICTs) were expected to contribute to foster democratic governance in at least five dimensions. I suggest briefly to frame a general discussion on the potential contribution and risks of artificial intelligence. ICTs were expected to increase 1) Democracy and democratization; 2) Transparency and accountability; 3) Citizen participation and, as a consequence, democratic legitimacy (top-down approach); 4) Civic empowerment (bottom-up approach) and 5) Efficiency in public policy implementation.

First, the spread of the internet and digital media was expected to revitalize democracy and promote democratization in authoritarian countries. It did not happen or at least not in a linear way, because structures and political context matter and to reach the goal is far more complex than based on just introducing ICTs. Second, ICTs were expected to promote transparency and access to information and it did happen to a considerable extent, feeding the fight against corruption. But the potential to foster transparency is limited by political variables and has not reached the expectancy; in fact, in times of COVID-19 the lack of transparency has increased in some areas. Third, many governments around the world and particularly in Europe and at the European Union level have launched platforms to promote civic deliberation and to collect public opinions. The most typical experience is one in which goals and mechanisms to process the contents generated are not clear, only a few participate and there is no connection with policy-making; quite often these initiatives look like just a symbolic platform to claim that the government is engaged in promoting citizens participation. At the level of cities, the picture could be better, but at the level of the states, these initiatives have offered a very limited outcome. Fourth, bottom-up initiatives to collaborate creating applications for improving urban life and or to control the authorities and or to intervene in policymaking are for sure a promising development. Access to raw data has also offered good opportunities for such developments, which connects with the fifth, policy implementation. ICTs have reinforced governments' capacities to provide better answers. The risks are related to reinforcing inequalities, given that the most active users of these applications are quite often better educated and living in wealthier areas than the average.

Can AI contribute to a policy-making more resilient? It can contribute, but other political factors and contextual factors matter and need to be taken into account. What can we learn from the digitization of policy-making processes? A top-down approach is less likely to succeed while collaborative forms of governance are likely to make a difference. Under what conditions could AI support civil society participation in policy-making? By providing information, but legitimacy, and public discussion require human intelligence and there is an evident risk of political manipulation which needs to be prevented.

Supporting youth participation in decision-making processes

Dr. Jasmine Lorenzini, University of Geneva

What we know about political participation is that young people tend to participate less in politics, and this is especially true for institutional politics. Yet if you take the example of protest or boycotting, outside of

the institutional realm of politics young people do participate, and even more than older citizens. This suggests that some young people want to participate, so can artificial intelligence enhance youth contribution to decision-making processes, in particular, institutional politics?

First, there are different barriers to youth political participation, and second there are multiple pathways to influence decision-making processes. Sidney Verba et al.[26] wrote about equal political participation and identified three barriers to political participation. The first one is, because people don't have the motivation to participate, also young people distrust elected representatives, public policies do not address issues that are important to them, and this is more the substantive bias in the democracy. Second, young people do not participate because they lack key resources that support political participation (political knowledge, understanding of the rules of the institutions, understanding where one can have an influence on the decision-making processes). An important resource is time, for instance, youth in precarious working situations have no time and no money to support their everyday expenses, but also to invest some time in politics. Political action requires a financial investment (joining a party or an association) and time. So some youth do not participate because they cannot, they don't have the resources to do that. And, finally, if nobody asked them to participate, even for young people who are motivated to participate and have the resources to participate if nobody asks, they might not participate in the end. Focused on youth but it's very much the same if you think about women or foreigners, or any other vulnerable group.

During the last election of representatives in Switzerland, it was very much said the Friday's for future[27] movement but also the women's strike influenced the outcome of the election, with more young people in the parliament, more women, but also greener elected representatives. So, we have a recent example of this influence of contentious or protest politics on the election, processes, and outside of election period, can also have an influence because it can bring issues on the agenda and draw attention to these issues.

Artificial intelligence could play a role in increasing, for instance, the visibility of young candidates. As young people might suffer from the fact that they have fewer resources at the individual level but also at the collective level, although they would like to be elected and represent other young people, their chances of being elected, are much lower. And on the substantive representation, the idea could be to enhance the issues that are important to young people. So maybe analyze existing fora where young people meet and discuss about politics to identify problems they face, the difficulties they have, or more generally issues that they would like to see in the public debate to identify which public policies would better address the problems that they face, and the changes they would like to see happening in society. Of course, these way of using artificial intelligence is not only challenging in terms of how to do it, but poses, very important, ethical questions. Is it fair to use artificial intelligence to enhance the visibility of some candidates or specific social groups only? How this would or might reduce other inequalities, or create new ones?

AI and society: a digital inequality perspective

Dr. Moritz Büchi, University of Zurich

AI is part of the larger developments of digitization. As such, digital competencies of the population are to be further strengthened if we want people to contribute autonomously to political and social life[28]. There is a common perception that “everyone” uses the internet, but many digital inequalities remain, and in as far as AI is another step in the process of digitization, it will tend to repeat unequal access and outcomes. For instance, whereas more than 90% of the Swiss population use the internet, we still see extreme age

[26] See for instance Verba, S., & Nie, N. H. (1987). *Participation in America: Political democracy and social equality*. University of Chicago Press.

[27] https://fridaysforfuture.org

[28] See Büchi, M., & Vogler, F. (2017). Testing a digital inequality model for online political participation. *Socius*, *3*, 2378023117733903 and Carretero, S., Vuorikari, R., & Punie, Y. (2017). DigComp 2.1: The Digital Competence Framework for Citizens with eight proficiency levels and examples of use (JRC106281). Publications Office of the European Union. https://doi.org/10.2760/00963 (ePub)

gaps between younger and older adults in *mobile* internet use as a more recent technological innovation[29]. From a theoretical perspective on technology adoption and society as a communication system, acquiring new knowledge and skills is often proportional to what people already have and know, which exacerbates social inequalities over time[30]. When information flow or innovation increases – as is the case with AI developments – disadvantaged groups will keep falling behind and are asked by elites and politics to "catch up".

Digital inequality research generally discerns three levels – the first is having access to the digital technology in question, the second level is having the requisite usage skills and engaging in various activities afforded by the technology. The third level concerns specific outcomes such as finding relevant information[31]. Ultimately, sometimes within a lifetime or intergenerationally, outcomes feed back into the positions that individuals occupy in the social structure, which in the first place was what determined the access – social position markers like education are associated with access and use of digital innovations[32]. There is no reason to assume this basic mechanism of technology effects in society does not apply to AI. However, what AI does show clearly, as compared to for example the diffusion of smartphones, is that it is not just about a technical artifact; AI is a whole assemblage of knowledge and values[33], requiring a range of competencies to put its myriad applications to good use (as well as a social negotiation of what "good" use means).

If the governance of AI ultimately aims at quality of life[34], any effects of AI applications will be embedded in an existing social, cultural, and political structure and they will interact with all other social processes. If we increase the benefits through more widespread adoption of AI applications, the harms will also tend to increase: it is very difficult to selectively operate on this cycle and find policies that are perfectly suited to reduce harms without also reducing benefits. AI at various levels may very well increase efficiency and welfare, but these same technologies may lead to privacy breaches and have detrimental effect on some other aspects that we have decided to value as a society. The same way that books can support education or social media can support political participation (that is, not by themselves), there will be many undoubtedly beneficial effects of AI. Yet, as argued, the benefits may tend to advantage the already advantaged more, as has been shown with many technological developments. Implementations of AI to support civil society's participation in policy-making processes can thus only make sense as one of many tools, accompanied by initiatives working on long-standing social inequalities. Successful expansion of participation in policy-making cannot rely on technological solutionism and needs to consider socially unequal preconditions for and externalities of AI applications.

---

[29] see https://mediachange.ch/research/wip-ch-2019

[30] See Büchi, M. (2017). Digital inequalities: Differentiated Internet use and social implications [Doctoral dissertation, University of Zurich]. https://doi.org/10.5167/uzh-148989 and Rogers, E. M. (2003). Diffusion of Innovations (5th edition). Free Press ; and Tichenor, P. J., Donohue, G. A., & Olien, C. N. (1970). Mass Media Flow and Differential Growth in Knowledge. The Public Opinion Quarterly, 34(2), 159–170. https://www.jstor.org/stable/2747414; and BAKOM Bundesamt für Kommunikation. (2018). Strategie Digitale Schweiz. www.infosociety.ch

[31] See Van Deursen, A., & Helsper, E. J. (2015). The Third-Level Digital Divide: Who Benefits Most from Being Online? In L. Robinson, S. R. Cotten, J. Schulz, T. M. Hale, & A. Williams (Eds.), Studies in Media and Communications (Vol. 10, pp. 29–52). Emerald. https://doi.org/10.1108/S2050-206020150000010002 and Büchi, M., Festic, N., & Latzer, M. (2018). How social well-being is affected by digital inequalities. International Journal of Communication, 12, 3686–3706. http://ijoc.org/index.php/ijoc/article/view/8780

[32] Van Dijk, J. (2020). The digital divide. Polity.

[33] E.g., Cath, C. (2018). Governing artificial intelligence: Ethical, legal and technical opportunities and challenges. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376(2133), 20180080. https://doi.org/10.1098/rsta.2018.0080; and Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2; and Sujon, Z., & Dyer, H. T. (2020). Understanding the social in a digital age. New Media & Society, 22(7), 1125–1134. https://doi.org/10.1177/1461444820912531

[34] SEFRI Secrétariat d'État à la formation, à la recherche et à l'innovation. (2020). Intelligence artificielle. https://www.sbfi.admin.ch/sbfi/fr/home/bfi-politik/bfi-2021-2024/transversale-themen/digitalisierung-bfi/kuenstliche-intelligenz.html

How is AI used in Voting Advice Applications

Prof. Diego Garzia, University of Lausanne

Voting Advice Applications (hereafter: VAAs) are known to contribute to political knowledge and electoral participation. They are a nonpartisan online tool developed either by NGOs, universities, or combinations of them. The basic point of VAAs is to help their users casting an informed vote. The long-standing idea of civic voluntarism finds information and knowledge as preconditions for participation. Within the low information rationality framework, developed within cognitive psychology, we know that reasoning voters are expected to minimize information costs, and that the probability of them to participate is actually inversely proportional to the costs involved in gaining information. In this sense, existing research has shown that VAAs do provide cheap information that citizens find useful, and this increases knowledge of the political process and foster electoral participation as a result.

How does AI claim? There are four pillars in the making of a VAA. Basically, the long-standing technological features of VAAs lies with their algorithm, i.e., the calculation method at the basis of the provision of a voting advice. All the other steps are done by human intelligence. In fact, the questionnaire includes all the topics the researchers feel to be important to compare the parties and the voters. These issues need to be salient in the campaign in order to be relevant for the users.

When it comes to the selection of parties, the extent to which an ideologically unacceptable party should be included is still subject to normally human reasoning.

Other aspects of VAA making, however, have evolved in technological terms. For instance, on the one hand, we have the coding of party and candidates' positions. Until very recently, this was either done by means of actually asking the party or the candidate to fill the survey themselves to show where they were on issues, or by political scientists (based on expert's assessment, the analysis of the party Manifesto or the interviews and other written material). On the other hand, the calculation and display of the results have always been characterized by static a priori decisions. So, the makers will decide the shape of the political space of a given country, and force parties and users to adapt to the same political playground. More recent development is this idea of *dynamic scale validation*. The basic logic for this type of machine learning application is to exploit the early user data as a benchmark against which the ex-ante defined maps, could be evaluated. People are told to vote for a different party, based on who designs the map.

Another point of contention links to the idea that all issues matter the same in certain VAAs such as StemWijzer[35] of Wahl-O-Mat[36], whereas users may find that some policy issues are disproportionately more important than others. So, it has been proposed an adaptive algorithm that learns the configuration of the conceptual space from users' answers. AI can offer users a better understanding of the political competition that is closer to their own understanding of politics. To some extent, all of these applications are based on user's implicit expectations. But this could actually lower the educational, or simply informational role of VAAs. After all, the VAA phenomena is born out of an higher education background. And if we come with VAA that are closer to what people think already or is convinced already we may increase in precision, but we may decrease in educational terms.

We also found that in referendum, when the spectrum is too large, a topic like Brexit, which would encompass every type of ideological reasoning, every type of preexisting background attitude and even, sociological features, that would be way too complicated for a VAA to tell people towards a yes or no. And that would make even more noise, because the public was too uncertain and the issue was too complicated, but applied to Swiss initiatives, narrower and more specific, less political and more possible to disentangle the dimensions that we use in normal VAA questions.

[35] https://stemwijzer.nl
[36] https://www.bpb.de/politik/wahlen/wahl-o-mat/

3. AI & collective intelligence

Dr. Aljaz Kosmerlj: Consensus-Building Machines
Nathan Stern: Civic tech and Concertation: Lessons learned
Dr. Maxime Stauffer: Computational simulations to explore complex policy scenarios
Vanessa Hanifa: Foresight thinking for policy processes

Consensus-Building Machines

Aljaz Kosmerlj, Josef Stefan Institute

Artificial intelligence (AI), as it is used today, is first and foremost a tool which functions for specific tasks. For example, if an algorithm is tuned to identifying, if there's a cat in a picture, that algorithm will only do that particular task. It will not recognize cars in pictures, play chess or find an optimal driving route in a city. To put it differently, the models do not generalize and are typically quite narrow in scope. In commercial use, the high-level task or purpose of the algorithms is to maximize profit. An obvious example is how Amazon uses recommendation systems that identify the most likely next purchase so that they can promote it to you and incentivize more sales. In large-scale deployments these systems, though targeting these narrow commercial goals, can have unforeseen effects. This can happen through misuse, such as how C. Analytica used Facebook's add targeting system, arguably designed for less-nefarious traditional marketing purposes, to influence election results.

However, negative consequences can also be emergent from the system themselves. For example, the YouTube video recommendation service has been shown to promote conspiracy theory videos, because their users tend to spend a lot of time watching them. Since YouTube makes money by showing videos, the system is prioritizing those videos that increase chances of further watching. The company has no interest in the content of these videos.

From a governmental perspective, these tools could be used to build consensus in society. A lot of social media divides people because it connects you to people who share the same belief on one side and shows you opposing views in an inflammatory way on the other - anything that makes you engage. Again, the goal of these companies is not to make us angry, it's just that this makes us spend more time on their sites, which they can monetize. Unfortunately, this acts as a constant drip of discord into society. The same AI tools can technically be used in a way that fosters agreement. Methods such as clustering can find commonalities in people's views. By analyzing opinion points of crowds, we can identify areas of consensus and use that to guide public policy formulation. When building tools for this purpose, human-computer interaction design can help us carefully craft the interface so that we limit trolling in the interactions.

Governments should invest in such AI systems that have positive effects. The applications themselves need not be complex. For example, in order to improve transparency, all the discussion in the Parliament and other official governmental bodies could be automatically transcribed with speech to text engines. By parsing the transcriptions keywords and topics can be identified and notifications can be forwarded to interested individuals and organizations. The whole process is robust, can be fully automatic and requires (relatively) little infrastructure while directly improving transparency in democracy.

A more advanced example is the Sunflower movement in Taiwan. They built a novel online/offline consultation process they call Virtual Taiwan (VT). It is built on a set of open source principles and its core idea is to build policy from the bottom up through public interaction. By using an open source platform called Polis, members of the public can submit statements on a topic of public policy. Users of the platform then vote on the statements, expressing their (dis)agreement. AI services analyze this data in real-time and promote the statements where people agree and display them in the platform interface. Over time people

have been shown to converge in consensus areas which are then used to base policy on. The system has addressed issues such as regulation of Uber and electric vehicles in Taiwan. There are still open questions about transparency, fairness, robustness, and representation when using such a system on a larger scale, but results indicate that the approach can work.

Civic tech and concertation - Lessons learned

Nathan Stern, Peuplade.fr

We all know that tech has dramatically boosted the progress of the most diverse fields of activity: chemistry, logistics, medicine, animation, chess, etc. But what about politics? Is it better today? Of course not. Three factors may explain this relative failure.

First, apart from the consultation platforms launched by local authorities on very specific development projects, civic tech exceptionally manages to influence political decisions, regardless of their form: online petitions, collaborative ideation platforms, platforms for the collective elaboration of reforms or decisions, debate platforms, etc. One could question the institutions, that definitely have difficulty in welcoming citizen expression. But we can also question the quality of the contents produced on Civic tech platforms. This “quality-issue” is such that many contributions seem to be read only by their authors. Even the contributors don't always read each other on those platforms.

The second factor may be that civic tech platforms often offer an online transposition or traditional political dynamics: it’s all about referendums, proposals, debates, votes, quantitative logic (with the “one man, one vote” principle) In most civic tech platforms, citizens are invited to express themselves, propose ideas, debate and vote. Apparently, those actions seem to be the basis of any political involvement. But does it make sense to vote when nothing is done to get a representative audience? Is it useful to debate when the only resistance we encounter is that of people who disagree with us, and never the only resistance that really matters: the resistance of reality !

Citizens work like politicians. They refer to doctrines, principles, values, sometimes observations… but bringing factual data into the debate, relating to reality, local or foreign examples at the right level, this almost never happen. And most of all, reforms are not rated according to their supposed impacts but according to their political meaning and its ability to get interest. Politics – the amateur one as well as the established one – seems seen as a space for expressing one's will and not as a space in which one seeks solutions to problems. In a way, civic tech is currently inspired by politics as it is done. And its issues are the same than those of politics - except that citizens who participate online don't have the power!

The third factor may be that civic technologies are essentially tools of expression - just like social networks. Civic tech seems to be very focused on citizen participation… just as platform’s editors, whose only focus is, first and foremost, participation, and not the quality of the contributions. The "Citizen participation" seems to be taken as an end in itself. In this context, the very notion of quality is almost taboo: Both elected officials and platform editors are afraid to position themselves as a judge. They want to show their respect for the people's word, which is sacred. Moreover, for the most part, on the functional level, civic tech platforms are inspired by social networks. They highlight the new contents, the most popular ones or those that create engagement. Contributors never question their opinions, they never express any doubts. They just say what needs to be done from their point of view... In short, they express themselves in this narcissistic omnipotence which is characteristic of social networks.

But a reform is not only an expression of the popular will. A reform is also a program of transformation of reality. We can consider programs to lower the number of COVID victims, of CO2 emissions or the number of bicycle thefts, etc. Strictly speaking, only programs that work should be "installed" in our country. But we don't know in advance what will work.

Fortunately, it is possible to estimate the impact of a reform without waiting for its implementation, especially if the reform has already been implemented in other contexts, and that the impacts in question can be observed, measured, and documented. If civic tech is to represent real progress, It will have to produce reforms of such quality and robustness that politicians will have no choice but to draw their reforms from it. But the possibility of progress is condemned if we don't feel legitimate to assess the human, economic, social or ecological impact of the proposed reforms. What If we consider political programs the way we consider computer programs? Their value only lies in their ability to bring the expected outcomes.

Computational simulations to explore complex policy scenarios

Dr. Maxime Stauffer, Geneva Science-Policy Interface.

Complexity challenges policymaking for different reasons and computational simulations can help navigate this complexity. Complex systems have many, interacting parts and are subject to negative and positive feedback loops that generate nonlinear dynamics. For instance, a small policy could have huge effects because they are amplified by feedback loops in the system, or a large policy could be completely absorbed by the system and have zero impact. These nonlinear dynamics challenge predictability and raise uncertainty. Therefore, guiding policy-making based on the expected consequences of policies is very difficult and depends on our ability to navigate complexity. Moreover, an additional problem is that the data we use to make decisions in complex systems usually account for macro variables, like GDP, COVID cases, etc. The nature of this data highly contrasts with the reality of policy-making, which often attempts to shape micro-level processes such as human behaviour. One way to solve this problem of nonlinearity and mismatch between available data and policies is to use computational simulations. Here are three examples[37]. For additional examples, see this paper by Dirk Helbing and colleagues[38].

One example is to support behavioral policies during pandemics. A study by Ayla and her colleagues[39] shows how they collected data on COVID-related sentiments and behavior based on social media. Collecting those are granular data at the individual level provides insights on how people actually feel and think about a complex situation. They use this data to understand different behavioural and sentimental phases related to the pandemic, thereby informing policies that take the reality of human behaviour into account. The second example is on prioritizing economic policies. In a network model based on empirical data, Guerrero & Axtell[40] found a power law about the relationship between firms and employment growth. In classical economics, we tend to assume that each firm is responsible for employment growth proportionally. In contrast, they find that 10% of firms are responsible for 90% of employment growth, thereby informing which firms policies should target. A third example pertains to conflict situations where decisions are sensitive to short term effects. A computational model by Bhavnani and colleagues[41] is able to predict violent events in the city of Jerusalem as a function of micro-level, social processes. Subsequently, they could run safe, what-if scenarios of policy proposals such as the Clinton plan or the Palestinian proposal to identify whether they lead to more or less events of violence.

All in all, my point with this short presentation is that the field of complexity science can: (1) provide granular insights related to individual and group behaviour to craft more effective, socially attuned policies; (2) draw insights from interaction topologies and their impacts on system behaviour, to increase policy

[37] See Complexity Lab. https://www.complexitylab.ch/post/how-complexity-science-can-bring-useful-insights-for-effective-policy-making

[38] Helbing, D., Brockmann, D., Chadefaux, T., Donnay, K., Blanke, U., Woolley-Meza, O., ... & Perc, M. (2015). Saving human lives: What complexity science and information systems can contribute. *Journal of statistical physics*, *158*(3), 735-781.

[39] Aiello, L. M., Quercia, D., Zhou, K., Constantinides, M., Šćepanović, S., & Joglekar, S. (2020). How epidemic psychology works on social media: evolution of responses to the covid-19 pandemic. *arXiv preprint arXiv:2007.13169*.

[40] Guerrero, O. A., & Axtell, R. L. (2013). Employment growth through labor flow networks. *PloS one*, *8*(5), e60808.

[41] Bhavnani, R., Donnay, K., Miodownik, D., Mor, M., & Helbing, D. (2014). Group segregation and urban violence. American Journal of Political Science, 58(1), 226-245.

cost-effectiveness; and (3) explore possible futures in highly charged contexts, providing evidence-drive assessments to mitigate risk and reduce violence.

Foresight thinking for policy processes

Vanessa Hanifa, Haute École de Gestion de Genève, HES-SO

Foresight is kind of a way of thinking more than a discipline per say. Foresight is a structured and explicit exploration of multiple plausible futures in order to inform and support decision making, right now. The futures explored provide organizations with new pairs of glasses to look at the present, pay attention to weak signals out of their radars and enlarge their mental models. Foresight connects with many disciplines, such as management and social sciences, as well as arts & design given their very nature and ability to bring out feelings and intuitions about the future.

Organizations look at situations as being complicated, they do analogies, extrapolations, probabilities and statistics about the future. But when you do foresight you acknowledge that the world is complex, and not just complicated. So that basically means in terms of public policy that you cannot just go through the regular planning processes, you have to go beyond the limited and limiting relationship that public policymakers have with the future (driven by probability and desire). To do so, public policymakers need to explore possible futures. Foresight in public policies has three major impacts directly related to decision-making.

- First: anticipate or reframe new opportunities and challenges and putting them on the agenda.
- Second: develop new thinking about what are the best policies to address those opportunities and challenges.
- Third: evaluate whether proposed strategies are future proof against a set of future scenarios.

What about foresight and AI? The use of foresight methods is highly relevant to establish & structure an open-minded dialogue on AI, its different applications, the opportunities and threats, beyond the “official narrative” about the future and beyond hopes and fears. AI can provide valuable tools to support the detection of weak signals, which are at the heart of foresight methodologies and whose detection is key for a more resilient society.[42]

[42] Gauthier, T., & Hanifa, V. (2020). *Prospective* (No. BOOK). EPFL Press.

4. AI & The Political

Prof. Solange Ghernaouti: AI and power asymmetries: challenges for civil society and policy making
Prof. Anna Leander: Redesigning the political with blockchain
Abhimanyu George Jain: Subjects and politics in civil society's turn to AI

Artificial intelligence and power asymmetries: challenges for civil society and policy making processes

Prof. Solange Ghernaouti – UNIL, Swiss Cybersecurity Advisory & Research Group, Fondation SGH – Institut de recherche cybermonde.

What kind of rationality determine AI in policy making processes? Who are in charge? To answer it, we have to recall some fundamentals about digitalization of our society. Initiated in the last century, the digitalization of all activities is increasing and policy making processes are no exception. It raises several questions as for example, how can we ensure the autonomy of decision in policy making? And how data and artificial intelligence can modify and influence society's choice? We have already the case of information manipulation, a kind semantic war that occurs at national and international levels. Fake news are a reality, they could originate from local actors or foreign one, including state actors. The datification and the plateformization of the world has been developed from a unique model. That of the capture of data given by users and the exploitation of them by hegemonic actors.These private actors have been able to deploy infrastructures, tools and services on which society has become dependent. In doing so, they have developed capacities for collecting, processing and storing data, which gives them a leading role in the field of big data analysis and artificial intelligence. Real opportunities exist around the use of data, but dependence on infrastructure and service providers generates new forms of expression of power. It is an asymmetric power between those who develop, own, manage infrastructures and impose new practices and new ways of thinking, doing and existing, and those who depend on them. Users often have no choice but to subscribe to the conditions imposed by the strongest actors and to submit to their unilaterally defined rules. The capture and exploitation of data is part of a political, economic and techno-ideological project, that is part of a logic of power, a power without limits, for global and unlimited market.

Moreover, the digital economy allows the development of the economy of surveillance. It rises different problems that impact policy making processes and the way that public policies are designed. Monitoring the population, logging each action, movement, behavior, even feeling. It permits to have a perfect and detailed image of the society, to know trends, authorize any kind of the prediction and so of manipulation. Some rewards (nudges) or penalties can be assigned in real time to manage the way people acts. All these can modify the nature and the essence of government in an unprecedent way.

The government and public entities are becoming more and more dependent of private actors, mostly foreign actors, who own all the data and mechanism involved in. That increase power asymmetry and the gap between digital providers and civil society. We should also not forget the capacities of malevolent actors to destabilized and gain power by deploying cyberattack against infrastructures, should it be for criminal, terrorist or conflicts motivation. All have a huge impact the democracy functioning.

The industrial and commercial world proceeds by defining standards, by defining solutions that fit all. According to an industrial and commercial ideal, every individual, by extension the mass of people, should be obedient and predictable consumers. According to a security ideal, any individual should be a docile citizen. In both cases, it is a matter of dealing with an ideal clone of an individual based on a human mathematical model optimized by algorithms. It's not about diversity and plurality. It is about dealing with uniformity by computerized standardization that fit an economic logic. An algorithm isn't neutral, it serves a certain vision of rationality, of our society. If the logic of economic optimisation and techno-economic rationality is pushed to the extreme, we can end up with a kind of eugenics of thought and behaviour.

Artificial intelligence could lead to a kind of global unification, to a sanitization of the society. Today, most of artificial intelligence technologies are in charge of measuring in real time our actions and checking their adequacy to a normative framework of a predefined model. In this context, if the objectives of artificial intelligence are to establish an organisation of society according to utilitarian criteria, if the computerization of society is based on rationalization by focusing on certain factors that affect diversity, If so then, it is not surprising that this is reflected in artificial intelligence's algorithms.

Can we honestly answer the following question : How can we ensure that systems can learn (machine learning) by freeing themselves from initial biases, bias in the purpose and objectives of the algorithms?

Asymmetry between "human in the loop" and "human out of the loop" is important and will be more and more critical. Some aggravating factors are related to the development time and the time to market, which impose new practices, without it being possible to establish quality, security and safety controls. In fact, there is no time, no evaluation or certification process between invention, design and market adoption. The difficulty is increased by the context of competitiveness and competition between market players who want to be the first in the market and an immediate return on investment.

We have several opportunities to promote the development of and to design secure, open, transparent, and inclusive artificial intelligent systems. That means that systems should be transparent and "open source". This will make them more easily available to developing countries and reduce the digital gap. If open data are used to develop and train artificial intelligence system, so these systems should benefit to the community, concept of open artificial intelligence. Certification of products and services incorporating artificial intelligence could be then possible.

Being able to check, audit, test and certify products in order to use them with full knowledge of the data used and the computation perform, will allow us to choose or to renounce to some products. Confronting the new forms of bias generated by digital technology and artificial intelligence requires making them visible and questioning the social and political conditions for innovation. Confronting the new forms of violence generated by digital technology and artificial intelligence in particular, requires a strong political will and no doubt recourse to regulation, because ethic and the free will of actors will not be enough. Assigning responsibilities is mandatory!

Data is an instrument of power practice. Data and computation are at the heart of power activities. Concentration of data is synonym of concentration of wealth, wo is in the hands of only a few actors. It allows the increase in knowledge, in controls and the expansion of surveillance. This correspond to the willingness to be able to have more power, to keep it and to set up and to guide the way of civil society behave. It can destroy any form of dissidence.

In this context of digital economy, of rationality and data-oriented society models: we should consider the following questions: Who produces them, who finances them, who exports them? What kind of counter-power can exist? Is separation of power can be effective? What is their legitimacy and level of acceptability from a civil society perspective? We could not answer the question ": What are the mains skates and challenges for civil society regarding artificial intelligence and policy making processes?" without overlooking the following crucial issue: Could AI contribute to the end of politics? For me, this is the biggest challenges we have to consider.

Redesigning the political with blockchain

Prof. Anna Leander, Graduate Institute of International and Development Studies, Geneva.

Blockchain has been extensively used in politics and there has been a lot of enthusiasm, around blockchains to loop solutions to political problems and particularly to the problem of trust in political processes. However, the most significant potential of blockchain resides in its potential to *redesign* political processes, subjects, subjectivities and atmospheres.

Blockchain is shorthand for distributed ledger technologies that distribute information through AI so there's the link to AI into different distributed spaces so that it becomes difficult to hack and tamper with. A lot of the politics is located actually in the design of the technology of the blockchain. And it's that protocol that become the core of the political process. In 2013, in elections in Santiago de Chile but also in referendum in 2016 in Colombia, blockchain was used concretely to allow people to participate. A lot of companies have been selling blockchain solutions, smart Matic[43] is the biggest one

Blockchain is heralded as a 'trust technology' that can be used to ensure trust in electoral processes. However, far more fundamentally Blockchain is redesigning all aspects of citizen participation and political space. First, by redesigning what the political subject is. It allows you to raise political subjects that weren't already there before or change their status as shown e.g. by the blockchain Robin Hood cooperative[44]. Second, blockchain redesigns the topology, what political topics. One example is the politics of one of biggest cryptocurrency's initiatives in Africa, the Kenyan Sara foo initiative, used to reshuffle the politics of finance by creating credit for people that would otherwise not have access to it. Finally, the is redesigning the atmospheres of politics, revining some of the spirit and enthusiasm that surrounded Steven Barlow's manifesto of independence for cyberspace.

As Hito Steyerl[45] has suggested we might do well to think about blockchain less as a currency and more as an under-current that runs under and has the potential of invigorating gives the infrastructure to democracy and political process. Blockchain alters — redesigns — the what who and how of politics. It reframes not only citizen participation but also what a political citizen is.

Politics and subjects in civil society's turn to AI[46]

Abhimanyu George Jain, Graduate Institute of International and Development Studies, Geneva

Today I would like to center the fate of politics and of the subject[47] in the broader discussion of the utility and benefits of AI for civil society through the example of a Brazilian organization called Operação Serenata de Amor (OSA) has developed an open-source AI robot called Rosie, which scrutinizes receipts submitted for reimbursement by Brazilian MPs for fraud.[48] The large number of receipts makes manual scrutiny of all receipts impossible. Enter Rosie, who examines the receipts which are too small to merit investigation by the government and flags suspicions for further investigation. I have certain concerns about this model, but I want to be very careful to emphasize that these concerns are not intended as criticism. I think OSA's work is marvelous and creative, and I raise these concerns in relation to the underlying logic of the Rosie solution, which also seems to represent the bounded imaginary of civil society's conceptualization of the uses of AI.

One problem here is that the logic of Rosie suggests an abandonment of efforts to tackle corruption as a social and political problem. The objective is not to prevent corruption. It is to make corruption unconcealable and thereby to eliminate it. The ends may be the same or at least similar, but the means are very different. The difference at play here is the difference between political and technocratic solutions. The logic of Rosie suggests the rejection of the idea of corruption as a socio-political problem to be solved and prevented, in favor of an idea of corruption as an unavoidable but undesirable social fact to be

[43] See https://www.smartmatic.com
[44] See https://www.robinhoodcoop.org
[45] Steyerl, H. (2014). Proxy Politics: Signal and Noise. *e-flux journal*, *60*.
[46] A more detailed version of this presentation is available at <https://lawsandwarcrimesdotcom.files.wordpress.com/2020/12/politics-and-subjects-in-civil-societys-turn-to-ai-vf-201209-website-1.pdf> accessed 20 January 2021.
[47] The understanding of 'politics' and 'subject' which inform this analysis are drawn respectively from Chantal Mouffe, *On the Political* (Routledge 2005) and Rosi Braidotti, *The Posthuman* (Polity 2013). The arguments made here draw on Mark Andrejevic, *Automated Media* (Routledge 2020).
[48] 'Operação Serenata de Amor' <https://serenata.ai> accessed 20 January 2021.

eliminated. There is a shift from prevention and intervention at the causal level, to blanket pre-emption and a disregarding of causal and contextual factors.

The prioritization of technology over politics in the Rosie model is also reflected in the implied attitude towards the objects of regulation – the MPs. In the absence of a technocratic solution, the regulation of this corruption would work on the logic of deterrence. The idea would be that the fear of being caught would drive compliance with the law, and in this way the requirements of the norm would be internalized. Rosie is unconcerned with inducing compliance: the locus of operation shifts from the MP to the MP's environment; the operation shifts from inducing compliance to making non-compliance impossible. Implicit in this shift is the assumption of the inalterability of the MP's inherent corruption, and the denial of the subjectivity of the MP. The environment displaces the MP as the object of regulation.

Let me try and push this analysis further, to try and understand why and how civil society's turn to AI, insofar as Rosie is emblematic of the broader possibilities, entails these risks to politics and subjects. I think that at least part of the answer lies in limited imagination. Rosie's innovation lies in being better than humans at doing what humans were previously doing. The human version of this response – expense claim scrutiny as the answer to corruption – was limited by human capacity and so was forced to engage with the problem on a political level. But with Rosie, every expense claim can be scrutinized, and cheap, effective and exhaustive expense claim scrutiny can become the sole focus of anti-corruption efforts in this context, triggering the concern for politics and for subjects that we have discussed. In effect, technology does not only change what *can* be done, it inevitably changes what *should* be done. Rosie does not only facilitate expense claim scrutiny, she engineers a shift in the policy for responding to this form of corruption.

There are two aspects of this shift from 'can' to 'should' that I want to highlight before concluding. First, the shift is not a problem in itself. It is a problem only if it goes unnoticed and unanalyzed. Second, acknowledging the shift from 'can' to 'should' unveils glorious disruptive possibilities. Recognizing the possibility of using technology to change socio-political contexts opens the door to deploying the potential of AI in radical and creative ways. Rosie need not be restricted to outperforming humans at the only tasks that humans could devise.

5. AI for Governments

Olivier Leclère: Chatbot and political participation in Geneva
Jamie Berryhill: AI uses and considerations for governments and partners

Chatbot and political participation in Geneva

Olivier Leclère, Canton of Geneva

In 2017, with some colleagues, we developed a chatbot called LandmanBot[49]; it had two goals. The first one was to help Geneva's citizens to cast their vote through the online e-voting platform called CHvote[50]. The bot could answer questions like "until when can I cast my vote". The second goal was to provide the results of the elections and referendums. LandmanBot has been used by the canton of Geneva during one year, for answering questions about four referendums and two elections.

When you develop a chatbot, you have many choices to do. First, you have to choose which medium is the most suitable for your end users. You can either use an existing app to chat or develop one or integrate a chat in your website. For this project, we chose Facebook Messenger. At that time, Messenger was widely used, thanks to Facebook, and so the end users did not have to install another app nor go to a specific website, in order to discuss with LandmanBot. You then have to choose a solution to process the natural language (NLP). You can use a cloud solution, or you can develop your own. Due to data privacy for political rights, for the first part of this project, we decided to develop our own solution. We wanted to control the whole process and we wanted to know where the data was processed.

We started by developing the helpdesk part, in order to help people to cast their vote online. It was the toughest part, because people wanted to discuss with the chatbot, as if they were discussing with friends. They were sending short messages and emojis. We then had to feed LandmanBot with answers, because most of the important information upon political rights are only available in PDF files, which are not structured data. We also found out that in the case of elections, people wanted to know whom to vote or if they should accept or refuse the referendum. As an authority, you can never give answers to those questions. Then we finally developed a service to give the results and notify people when they were available. In this project, it was the most used functionality.

Using Facebook or a cloud solution to process data is against the law on protection of personal data for the public sector in Geneva (LIPAD). Due to this restriction, LandmanBot has never been the main channel for contacting the administration, it was an alternative channel.

Developing a chatbot, in order to increase the participation in the civil society, is difficult for a government, but the government should open its data (the State of Geneva offer all the informations about candidates and results as open data), in order to facilitate the development of new services. Using AI for a chatbot is important. It will facilitate the processing and the understanding of the questions. AI could also help to increase the quality of the service, for example by knowing when to send notifications / reminders.

Developing chatbots will facilitate the way people interact with an enterprise or a public service. A chatbot can work 7 days a week and 24 hours a day. It can handle, thanks to AI, every day more complex questions, but it should always allow human being to take control, to give a more precise answer. Let us hope it will increase the political participation. AI and chatbots are tools able to help us to deal with our daily tasks, and to remind us for the less regular one.

[49] See https://www.facebook.com/LandmanBot/
[50] See https://chvote2.gitlab.io

AI uses and considerations for governments and partners

Jamie Berryhill, Organisation for Economic Co-operation and Development (OECD).

Hello World! artificial intelligence and its use in the public sector[51] is a report on what to consider when approaching the topic of AI. One third is explainer one third a thorough review of the public sector landscape, one third, deep dive into government implications and considerations when government are exploring AI. So how the public sector itself can transform its own processes and services and engagement mechanisms. We found that some of the issues that were most prevalent, and we tried to provide advice and examples to this report is just one is evaluating options we saw a lot of countries that were just trying to do AI for AI's sake. There was not very much public sector-focused problem definition or exploration of what actual needs are and how those needs could be addressed with AI, or some alternative solution. There's a big focus on just everybody else is doing this, so we need to find more ways to use AI.

AI ethics has been another major point of discussion. OECD has its own AI principles, and those are fairly high level. Some governments have created more granular ways to implement ethical principles. For instance, Canada did an algorithmic impact assessment that helps identify actions needed to mitigate risks. In addition, the report covers access and use of data, trying to cover different aspects where governments are both opening their own data. Finally, a big focus is securing internal, external talent so new methods of providing vetted vendor lists for procurements finding new ways to upskill and train existing staff, engage in different partnerships with civil society or businesses.

Regarding AI and civil society, the citizen lab in Belgium is a great example of building a public consultation platform in take on citizen views and opinions at a scale, and synthesize the results and understand things about different, helping civil servants and government leaders make decisions about public policies and services and implement. Also, in Australia in a project called Converlens[52] is a kind of public consultations being able to receive a lot of input and feedback also from social media networks. We see different levels of maturity needed to really succeed when exploring and using AI. When it comes to digital government in general, and we have identified six core factors that we think make up a mature digital government: one digital by design. The second one is government as a platform, a third one being a data driven public sector, another factor being open by default and then finally just user driven and proactiveness. These are classified in the OECD's new Digital Government Policy Framework.[53]

[51] See https://oe.cd/helloworld
[52] See https://converlens.com
[53] See https://www.oecd.org/governance/the-oecd-digital-government-policy-framework-f64fed2a-en.htm.

List of Participants

Jamie Berryhill: is a Public Sector Innovation Analyst at OECD since 2016. Prior to joining OECD, he was Chief of policy, Budget and Communications at the White House Office of Management and Budget (OMB) from 2014 to 2016 and Senior Analyst on the Forensic Audits and Investigative Service (FAIS) team at GAO from 2009 to 2014.

Dr. Moritz Büchi: has been a Senior Research and Teaching Associate in the Media Change & Innovation Division since January 2017. He held a visiting position at the London School of Economics and taught at the University of St. Gallen. He is the coordinator of the World Internet Project surveys in Switzerland. In 2019, Moritz Büchi was awarded a DSI Fellowship and a Research Talent Development Fund grant. His current research focuses on digital well-being and overuse, privacy, and computational methodology.

Dr. Alexandre Dormeier Freire: is a sociologist of education, lecturer and researcher at the Graduate Institute of International and Development Studies where he also holds the position of Director of the Executive Development Policies and Practices programmes (DPP). He earned a PhD in Development Studies of the Graduate Institute. He is currently the lead researcher of a project funded by the MENA Learning House of the Swiss State Secretariat for Education, Research, Training and Innovation, which examines alternative social innovations mitigating violence in the Occupied Palestinian Territory. His research focuses on the links between children, families and schools and how these links can be harnessed to increase children's resilience to violence in contexts of protracted crisis. In recent years, he has also conducted various research projects on the transformations of contemporary education systems, focusing on children/school/family relations, the internationalization of education, social inequalities, education/employment relations, educational choices and school dropouts.

Dr. Jérôme Duberry: is a Research Fellow at the Dusan Sidjanski Centre of Excellence in European Studies, Global Studies Institute, University of Geneva, and Research Associate at the Albert Hirschman Centre on Democracy, Graduate Institute. His research activities focus on the participation of civil society in policy-making processes, and more particularly on their use of digital technologies, including artificial intelligence. His research interests extend to the different forms of governance in Europe, and in particular the federalist project. In 2019, Jérôme obtained a grant from the Swiss National Science Foundation (SNSF) to explore how artificial intelligence can contribute to citizen participation. Jérôme holds a PhD in International Relations from the University Pompeu Fabra (Barcelona, Spain).

Prof. Diego Garzia: is an SNSF Eccellenza Professor of Political Science at the University of Lausanne, and also a recurring Visiting Fellow at the Robert Schuman Centre for Advanced Studies in Fiesole. He holds a PhD from the University of Siena. His ongoing research project titled “The Rise of Negative Voting” is funded by an SNSF Eccellenza Professorial Fellowship of the Swiss National Science Foundation (2020-2024). With his team, he investigates the extent to which an increasingly confrontational style of campaigning and political communication—and the rise of negative identifications in a context of partisan dealignment and strong political personalization—could all be driving a distinctive form of negatively-driven voting behavior.

Prof. Solange Ghernaouti: holds a PhD in Computer science and Telecommunications at the Pierre and Marie Curie University. She is a professor at the University of Lausanne and an international expert on cybersecurity and cyberdefence. She regularly collaborates with various United Nations, European and government institutions as well as with private corporations. She leads the Swiss Cybersecurity Advisory and Research Group (SCARG)] that she has created and the research group in Complexity Science at the university. She developed an innovative interdisciplinary approach of digital security serving individuals, organizations and states at strategic and operational levels.

Sabrya Hamidi: holds a bachelor's degree in Sociology with a specialisation in Political Science from the University of Paris 8 and a master's degree in European Studies from the Global Studies institute of the University of Geneva. After working as a research assistant to Professor Sidjanski at the European Centre of Culture (2016), she is now a research assistant at the Dusan Sidjanski Centre of Excellence in European Studies (2020). She is also assisting Dr Jérôme Duberry at the Global Studies Institute in his research on AI and democracy.

Vanessa Hanifa: is research and teaching associate at the Geneva School of Management and Head ad interim of the foresight orientation of the Master of Science in Business Administration HES-SO. She holds a master's degree in engineering from AgroParisTech. She was Business Unit Manager of a consulting firm in strategy, innovation and exploration of unknown lands, where she accompanied institutional clients, companies and start-ups from all over the world. Her teaching and research work focus on the contributions of foresight to strategy and innovation

Abhimanyu George Jain: is a PhD candidate at the Graduate Institute of International and Development Studies and a research associate at the SNF-funded LAWS & War Crimes research project led by Professor Paola Gaeta. He has been admitted to the practice of law in India and in England and Wales, and has previously worked as a management consultant, as a corporate lawyer, and most recently, as a member of the counsel team representing the government of India in investment arbitration disputes.

Dr. Jasmine Lorenzini: is a research fellow at the Institute of Citizenship Studies at the University of Geneva where she works of food activism, political participation, and social movement activism. In the framework of her Ambizione project she works on food consumption and political activism, in addition she co-leads a project on seniors' engagement for climate. She also worked on youth long-term unemployment and its impact on economic, social, and political inclusion during her PhD and on protest in times of crisis during her post-doctoral research.

Dr. Aljaz Kosmerlj: is currently working as a researcher in the Department of Artificial Intelligence, Józef Stefan Institute, Slovenia. He is Doctor in Computer science from the Faculty of Computer and Information science of the University of Ljubljana. As an artificial intelligence researcher, he focuses on machine learning and data mining. He has experience with projects from both academia as well as industry and collaborated on preparation of successful funding proposals, both on the EU (FACTLOG -H2020) and national level.

Prof. Anna Leander: is Professor of International Relations/Political Science at the Graduate institute. Since 2013, she has been Professor of International Relations at the Pontifical Catholic University of Rio de Janeiro, where she will continue to teach a few weeks per year. She previously held positions at Central European University in Budapest, the Copenhagen Peace Research Institute, the University of Southern Denmark, the Copenhagen Business School, the Hanse-Wissenschaftskolleg, the Collegio Carlo Alberto and the Judith Reppy Institute for Peace and Conflict Studies at Cornell University. Her research is in International Political Sociology. She has worked extensively with practice theoretical approaches and she has a longstanding interest in the commercialisation of military/security matters.

Olivier Leclère: has been working in various functions, for more than 15 years, in the field of political rights for the canton of Geneva. Among other things, he has worked on the design and development of the Geneva electronic voting platform CHVote. He has also worked on improving the publication of voting and electoral information as well as results on the web and in open data. In 2015, he joined the team at the Chancellery of the Canton of Geneva in charge of voting and election operations.

Dr. Maxime Stauffer: is Science-Policy Officer at the Geneva Science-Policy interface. He is the co-founder of the Social Complexity Lab Geneva and of Effective Altruism Geneva. Previously, he was an advisor to the European Commission's Joint Research Centre and a researcher at the Graduate Institute of International and Development Studies. His academic background is in international relations and complex systems

science. He conducts research on how to best translate scientific knowledge and methods into policy and on how to improve collective decision-making.

Nathan Stern: is a social designer, consultant and public speaker. He created different platforms such as Peuplade.fr (2003), a social media which gathers neighbours. He co-created HappyWeek with Claude Van Leeuwen in 2014, which is a game stimulating children to accomplish daily tasks (homeworks for instance). Also, he created the Common Good Factory to bring "factory alike" solutions to people and organizations.

Prof. Himanshu Verma: Himanshu Verma is an Assistant Professor in the Industrial Design Faculty at TU Delft (Netherlands) and is also a core member of the AI Transparency Institute (Switzerland). He holds a PhD in Computer Science from Swiss Federal Institute of Technology Lausanne (EPFL). With a background in Ubiquitous Computing and Social Cognition, his work focuses on peoples' socio-technical interactions and experiences with intelligent and AI-powered artifacts and environments. In recent years, his work has contributed to the emerging domain of Human-AI collaboration. In particular, he has been engaged in developing a sociological conception of AI where AI-powered technologies are more attuned to societal needs and expectations. Within the framework of AI Transparency Institute, he has co-organized the AI Governance Forum to foster multi-stakeholder and interdisciplinary collaborations around a responsible and fair development and governance of AI technologies.

Dr. Yanina Welp: is a Research Fellow at the Albert Hirschman Centre on Democracy. She is also editorial coordinator at Agenda Pública and founder member of Red de Politólogas. Between 2008 and 2018 she has been principal researcher at the Centre for Democracy Studies and co-director of the Zurich Latin American Centre (2016-2019), both at the University of Zurich. She finished her Habilitation in 2015 at the University of St.Gallen, with the venia legendi in Latin American Studies and she holds a PhD in Political and Social Sciences from the Pompeu FabraUniversity (Spain). Her main areas of study are the introduction and practices of mechanisms of direct and participatory democracy, and digital media and politics, i.e. 'democratic innovations'.