Global Attitudes Towards AI, Machine Learning & Automated Decision Making

*Implications for Involving Artificial Intelligence
in Public Service and Good Governance*

Lisa-Maria Neudert, Aleksi Knuutila, Philip N. Howard

7 October 2020

Global Attitudes Towards AI, Machine Learning & Automated Decision Making
Implications for Involving Artificial Intelligence in Public Service and Good Governance

Lisa-Maria Neudert, Aleksi Knuutila, Philip N. Howard

SUMMARY

How do people perceive the risks of using AI (Artificial Intelligence) and machine learning tools in decision making? Using survey data from a sample of 154,195 respondents in 142 countries, we analyze basic indicators of public perceptions about the potential harms and opportunities of involving AI in our personal affairs and public life. Public understanding of AI—its risks and benefits—varies greatly around the world.

1. There are regional and East-West divides in public attitudes towards AI-driven automated decision making, with worries that AI will be harmful running highest in North America (47%) and Latin America (49%), and notably smaller proportions of respondents in Southeast Asia (25%) and East Asia (11%) concerned that AI will be harmful.

2. Enthusiasm and optimism around the potential of AI in decision making runs highest in China, where only a small proportion of respondents believe that the development of intelligent machines or robots that can think and make decisions in the next twenty years will mostly cause harm (9%).

3. Across the professions, business and government executives (47%) and other professionals (44%) are among the most enthusiastic about AI decision making, whereas workers in manufacturing (35%) and service workers (35%) are less confident that automated decision making will mostly help society.

Our findings suggest that putting AI to work for good governance will be a two-fold challenge. Involving AI and machine learning systems in public administration is going to require inclusive design, informed procurement, purposeful implementation and persistent accountability. Additionally, it will require convincing citizens in many countries around the world that the benefits of using AI in public agencies outweighs the risks.

INTRODUCTION

Artificial intelligence and machine learning have a growing role in public policy making and governance around the world. Moreover, there is an expanding gulf between how artificial intelligence and algorithmic systems work and how people say they work – especially when government tries to use AI. Significant violations of privacy, dramatic failures of machine learning, and persistent data bias have plagued efforts to involve machine learning systems in government decision making in the past. COVID-19 has supercharged the adoption of artificial intelligence technologies in the public sector, exacerbating challenges around the use of AI decision making for good governance.[1]

As a result, the risks of AI have emerged as a prominent issue on the global public agenda. Once imagined as a catalyst for efficiency and prosperity, experts are increasingly concerned about opaque automated decision making, algorithmic bias and short-sighted technical systems.[2], [3] But what are the nuances of public opinion about the AI decision making, and its role in private and public life? Frequently referred to as a "techlash"—a backlash against technology—fears surrounding automated information systems, artificial intelligence and machine learning have become prevalent among citizens in many countries of the world.

Experts have always pointed to potential ethical, socio-political, and economic risks associated with new information technologies including artificial intelligence—at the same time recognizing the opportunities and transformative potential of AI and machine learning for virtually any sector. Increasingly, public sentiments about many kinds of information technologies have grown more skeptical. This may be a result of a series of revelations about the bias, abuse and privacy implications of popular technologies in recent years.

People around the world perceive different risks in the technology they use. AI-driven decision-systems have shown to be biased to discriminate against Black people.[4], [5] Privacy violating coronavirus apps and digital interference in elections may leave many citizens fearful of the risks in

relying on information infrastructure.[6] As COVID-19 has become a public health emergency, the use of AI and digital technologies to combat the spread of the disease and develop vaccines have been the topic of much debate. The debate itself, however, has clearly been degraded by the spread of misinformation over the automated algorithmic systems of social media companies.[7] What is more, AI decision making often remains opaque with algorithmic "black boxes" obscuring the rationales leading to a given result.[8]

Issues surrounding the risks of technology have certainly featured prominently on the public agenda for many countries. Rarely, however, can we examine trends in risk perceptions in a global and comparative context. In this briefing note, we conduct a comparative analysis of perceptions on AI based on data from a large-scale survey.

DATA AND METHOD

Using data from the _2019 World Risk Poll_, we analyze basic descriptive statistics about public opinion drawn from a sample of 154,195 respondents living in 142 countries. The survey design, participant recruitment and interviews were conducted by Lloyd's Register Foundation and Gallup. A detailed methodological description for the _2019 World Risk Poll_ is available on the Lloyd's Register Foundation's website .

As a survey instrument, the _2019 World Risk Poll_ employed an interview-based survey design using both face-to-face and telephone interviews. The survey questionnaire was designed by Gallup and Lloyd's Register Foundation in an iterative, multi-step process. The initial draft questionnaire was based on a review of relevant literature and expert interviews. This draft was then piloted and refined for two rounds in several countries. The final survey was comprised of ninety-five questions including background and demographic information.

The final survey included diverse questions on a wide array of risks, including issues such as food safety and physical safety at the workplace, as well as technology-related risks. Not every respondent was asked every question. For example,

only people who had reported using the internet were asked certain relevant follow up questions relating to the use of the internet. The questionnaire was translated into the major conversational languages of each country.

More than five thousand interviewers conducted interviews, having first participated in standard Gallup training on research ethics, fieldwork safety and interviewing techniques. At least 30% of face-to-face interviews and at least 15% of telephone interviews were validated through accompanied interviews, re-contacts or listening to recordings of interviews. Interviews for the *2019 World Risk Poll* were conducted between 8 May 2019 and 17 January 2020.

At least 1,000 respondents were surveyed in each country. All samples were probability based and nationally representative of the resident adult population—as defined in-country. The coverage area included the entire country and the sampling population represents the entire population aged fifteen and older.

For countries where population information was available, participant selection was based on probabilities proportional to population size. If no population information was available, random sampling was used. For face-to-face interviews, random route procedures were used to select households and participants were selected randomly within the households. For telephone interviews, telephone numbers were generated randomly, using a list-assisted random approach, or were selected randomly from a registered listing.

The *2019 World Risk Poll* does not publish response rates on a country-by-country basis but the median response for fifteen regions is available. Across these regions, the response rate ranged between 6% on the low end for North America and 80% for Middle and Western Africa.

The margin of error used in estimating the unknown population proportion "P" for the *2019 World Risk Poll* can be derived based on the following formula, where "n" is the number of respondents:

$$\text{Margin of Error} = 1.96 * \sqrt{P*(1-P)/n}$$

The margin of error for a sample size of 1,000 with P=0.5 will be 1.96 * $\sqrt{0.25/1000}$, or 3.1%, under the assumption of simple random sampling. A detailed table of 95% confidence interval half-widths for various sample sizes is available in the *2019 World Risk Poll* methods supplement.

Results were reported globally, regionally, and nationally. Survey answers were weighted to ensure that samples were nationally representative for each country. Accordingly, larger countries received more weight than smaller countries because of population size. In addition, population statistics were used to weigh survey data by gender, age, and where reliable information was available, education or socio-economic status. Additional information about national sample sizes, response rates, population weights, error margins, confidence intervals and other country-specific sampling details is available in the *2019 World Risk Poll* methodology appendix.

Our team did not participate in the design or fielding of the instrument itself but has conducted the statistical and secondary analysis needed to identify national trends and a global context for public opinion on the risks of AI in decision making. No new data was collected or used for this analysis. The results reported in this Working Paper make use of the weights calculated for the *2019 World Risk Poll*. When calculating averages for regions, the averages for individual countries in the region were weighted proportionally to the population size of respective countries.

FINDINGS

The *2019 World Risk Poll* included an important question that measures public attitudes towards the development of AI in the future. It asked whether "machines or robots that can think and make decisions, often known as artificial intelligence" will "mostly help or mostly harm people in this

in the next 20 years". Respondents could choose between "mostly harm", "mostly help" and "neither". Our team performed a descriptive statistical analysis to determine opinions on AI across different groups in the population. We compared sentiments on AI across gender, education, individual-level income, employment hours, attitude risk and other variables. We present findings on perceptions of AI across geographic regions and professions, as attitudes on AI exhibited important differences across groups for these variables.

Table 1 shows both the share of people who believe AI will be harmful as well as the share of those who believe it will be helpful. The data suggests that in many regions around the world public opinion is divided on the matter. For instance, in Europe 43% think AI will be harmful and 38% believe it will be helpful. This is a relatively small difference in comparison to Latin America & the Caribbean, where 49% are skeptical about AI and only 26% believe in its benefits. Conversely, 59% of respondents in East Asia reported believing that AI decision making would mostly help. There are some

Table 1. Global Risk Perception of AI Decision Making, by Region

Region	Mostly Harm	Mostly Help	Neither	Do not Know
Latin America & Caribbean	49	26	19	6
North America	47	41	12	0
Europe	43	38	15	5
Central Asia	34	36	17	13
Middle East	33	38	19	10
South Asia	33	31	17	19
Africa	31	41	16	12
Southeast Asia	25	37	21	17
East Asia	11	59	12	18

Note: Public perception of whether the development of machines or robots that can think and make decisions in the next 20 years will mostly cause harm or mostly help, by region. Based on sample of 154,195 respondents in 142 countries, with averages weighted by national population size before regional aggregation. Additional information about country-specific sampling details, including response rates, population weights, error margins, and confidence intervals, is available in the *2019 World Risk Poll* methods supplement (link).
Source: Authors' calculations based on survey data collected between 8 May 2019 and 17 January 2020 for the *2019 Word Risk Poll*.

interesting features to the structure of public opinion on AI. Public opinion within some regions is evenly split, such that the portion of respondents in Central Asia, the Middle East and South Asia who think AI decision making will be mostly harmful is roughly equivalent to the portion of respondents

who think it will be helpful. Moreover, respondents in North America, Europe, Latin America and the Caribbean are mostly likely to have an opinion on the question, whereas larger portions of respondents in other regions are more likely to say they do not have an answer to the question.

Table 2 reveals the differences in perception on AI among different professional groups. The professional groups that most strongly perceive AI as a threat are mainly construction and manufacturing workers and service workers, but also to a similar degree business owners. Over 40% of these groups see AI as mostly harmful. Manufacturing and the service industries are economic sectors that have a long history of automation and AI is likely to change and displace work in the future.[9] Nonetheless, just below 35% of manufacturing and service workers believe AI will be beneficial in the future.

The professional group that is most optimistic about AI includes executives in business or government, with 47% believing AI will mostly help. Positive attitudes seem to be shared across what could be called white-collar

Table 2. Global Risk Perception of AI Decision Making, by Profession

	Mostly Harm	Mostly Help	Neither	Do not Know
Construction or manufacturing worker	42	35	17	7
Hired service worker: e.g. maid, taxi driver or maintenance	41	35	17	7
Business owner with hired workers	40	37	15	8
Professional: e.g. doctor, lawyer, engineer, teacher	39	44	14	3
Clerical, sales or other office worker	39	43	14	3
Small scale trader or self-employed without hired workers	37	35	17	11
Executive or official in a business or the government	36	47	15	3
Farmer, fisherman or other agricultural worker	28	38	17	17

Note: Public perception of whether the development of machines or robots that can think and make decisions in the next 20 years will mostly cause harm or mostly help, by profession. Based on sample of 154,195 respondents in 142 countries, with averages weighted by national population size before aggregation in professional categories. Additional information about country-specific sampling details, including response rates, population weights, error margins, and confidence intervals, is available in the *2019 World Risk Poll* methods supplement (link).
Source: Authors' calculations based on survey data collected between 8 May 2019 and 17 January 2020 for the *2019 Word Risk Poll*.

Figure 1: Global Risk Perception of AI Decision Making
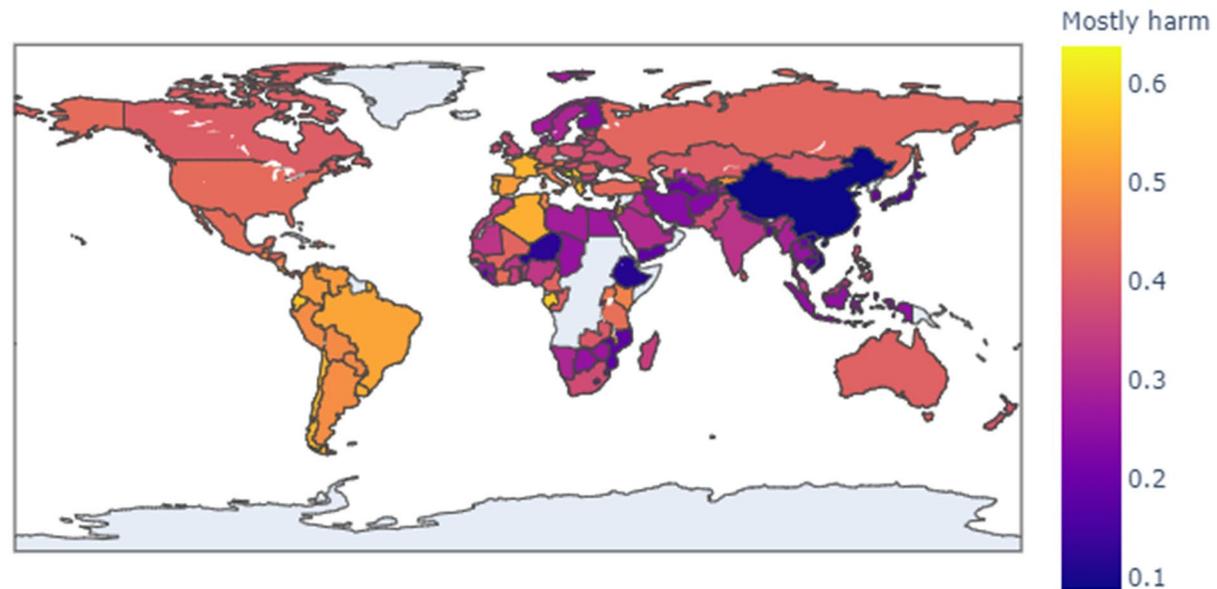
Note: Thermometer map of whether public perception of the development of machines or robots that can think and make decisions in the next 20 years will mostly cause harm. Based on sample of 154,195 respondents in 142 countries. Additional information about country-specific sampling details, including response rates, population weights, error margins, and confidence intervals, is available in the *2019 World Risk Poll* methods supplement (link). Country-specific figures can be more easily read from this interactive map (link).
Source: Authors' calculations based on survey data collected between 8 May 2019 and 17 January 2020 for the *2019 Word Risk Poll*.

professions, with office workers as well as professionals such as doctors or engineers having positive attitudes. Increasingly, agricultural workers are also much more likely to believe in the benefits of AI than to think that it is going to harmful. This group was more likely than others say that they had no opinion or to abstain from answering.

Figure 1 maps out public perception, country by country, of AI risks in decision making. In some countries most people believe that the development of AI will be mostly harmful, while in other countries a large majority think it will be mostly helpful. One outlier, for instance, is China, where only 9% of respondents believe AI will be mostly harmful, with 59% of respondents saying that AI will mostly be beneficial. Skepticism about AI is highest in the American continents, as both Northern and Latin American countries generally have at least 40% of their population believe that AI will be harmful. High levels of skepticism can be found in some countries in Europe. This skepticism runs high in Mediterranean countries such as Spain, Portugal, and Greece, but is highest in Belgium, where more than 50% of respondents expect the use of AI in decision making to be mostly harmful. Across all regions, when weighted by population size, 30% of respondents believe AI will mostly be harmful whereas 40% think it will mostly be helpful.

CONCLUSIONS

In this study, we made use of a unique, large-scale survey dataset to compare perceptions of technology-related risks between different regions, countries and professional groups. We find that, globally, people are concerned about risks associated with artificial intelligence.

Internationally, sentiments about technology are ambivalent at best. There are important differences between which risks are most prominent in a particular country. For instance, North Americans and people from Western Europe see the development of AI and robotics as more likely harmful as beneficial. Survey respondents in South and East Asia are much more likely to see these developments as beneficial.

The survey data available, though extensive in its geographic reach, only included a small number of questions about technology-related risk. More detailed research on the topic is required fully to describe the range of differences in the risks created by technology and how they are experienced by people in different positions. Further research would also be required to be able to explain these differences, and to see whether they relate to differences in technological environments or, for instance, cultural or social institutions.

Understanding public confidence in AI and machine learning is vital to the successful implementation of such systems in government. Each major misstep in AI will have consequences for public interest in having government agencies involve AI systems in governance. Even if the case for developing AI systems for good governance can be made based on administrative efficiencies, cost savings, or substantive outcomes, a government's implementation plan may be hampered by civic enthusiasm. Indeed, this evidence of public perception of the risks of AI in decision making suggests that public agencies will, in many countries around the world, struggle to convince citizens and voters that investing and implementing machine learning systems is worthwhile.

ACKNOWLEDGEMENTS

REFERENCES

[1]   P. N. Howard and Neudert, Lisa-Maria, 'AI Can Battle Coronavirus, But Privacy Shouldn't Be A Casulty', *TechCrunch*, 2020.

[2]   P. N. Howard, *Lie Machines: How to Save Democracy from Troll Armies, Deceitful Robots, Junk News Operations, and Political Operatives*. Yale University Press, 2020.

[3]   L.-M. Neudert and P. N. Howard, 'Four Principles for Integrating AI & Good Governance.', Oxford Commission on AI & Good Governance, Oxford, UK, 2020.

[4]   S. Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York, NY: NYU Press, 2018.

[5]   K. Lyons, 'Twitter is looking into why its photo preview appears to favor white faces over Black faces', *The Verge*, 2020.

[6]   J. Moreley, J. Cowls, M. Taddeo, and L. Floridi, 'Ethical Guidelines for COVID-19 Tracing Apps', Nature, 2020. [Online]. Available: https://www.nature.com/articles/d41586-020-01578-0.

[7]   A. Knuutila, A. Herasimenka, H. Au, J. Bright, and P. N. Howard, 'Covid-Related Misinformation on YouTube', Oxford Internet Institute, Oxford University, Oxford, UK, 2020.

[8]   F. A. Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge: Harvard University Press, 2015.

[9]   M. Arntz, T. Gregory, and U. Zierahn, 'The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis', OECD Publishing, Paris, 189. [Online]. Available: https://doi.org/10.1787/5jlz9h56dvq7-en.

ABOUT THE AUTHORS

Lisa-Maria Neudert is a DPhil candidate at the Oxford Internet Institute and a researcher with the Computational Propaganda Project, where her work is located at the nexus of political communication, technology studies, and governance. Her current research is looking into the public and private governance of policy issues surrounding disinformation through governments and social media platforms. She is the Co-Principal Investigator for the project's research on Europe supported by Civitates and the Secretary to the Oxford Commission on AI & Good Governance, OxCAIGG. Neudert holds an MSc in Social Science of the Internet from the University of Oxford, and a BA in Communication Science from the Ludwig-Maximilians-University in Munich. Selected as a Fulbright scholar, she previously studied at Georgetown University and the National University of Singapore. Neudert's commentary writing has appeared in several outlets including The Guardian, the MIT Tech Review and TechCrunch. Her research has been featured in several media outlets including the Washington Post, the Financial Times and Der Spiegel. She has presented her research at the European Commission, The Council of Europe and The World Economic Forum in Davos. Neudert is on Twitter @lmneudert.

Aleksi Knuutila is an anthropologist and a data scientist who studies new forms of political culture, communities and communication. He is interested in the application of interdisciplinary methods that combine interpretative, qualitative methods to the computational study of large-scale digital data. Knuutila works as a postdoctoral researcher as a part of the Oxford Martin Programme on Misinformation, Science and Media, as well as the Alternative News Networks project.
Knuutila completed his PhD at the Digital anthropology programme at University College London. For his thesis he undertook long-term ethnographic fieldwork with a contemporary monastic community in inner-city Austin, Texas, studying the connections between their communal practice and political imaginary. After completing his PhD, Knuutila was commissioned by the Finnish Cabinet office to build computational models for detecting hate speech on social media and used it to understand the structures of political groups disseminating hate speech. Knuutila has also worked in advocacy, in areas of digital rights, access to information and

political transparency. He administers the Finnish digital archive for Freedom of Information requests (tietopyynto.fi) and has applied FOI to make available the first datasets on meetings between lobbyists and Finnish parliamentarians.

Philip N. Howard is a professor and writer. He teaches at the University of Oxford and directs the Oxford Internet Institute. He writes about information politics and international affairs, and he is the author of ten books, including *The Managed Citizen*, *Pax Technica*, and *Computational Propaganda*. He has won multiple best book awards, and his research and commentary writing has been featured in the New York Times, Washington Post, and many international media outlets. Foreign Policy magazine named him a "Global Thinker" for 2018 and the National Democratic Institute awarded him their "Democracy Prize" for pioneering the social science of fake news. He has testified before the US Senate, UK House of Parliament, and European Commission on the causes and consequences of fake news and misinformation. His latest book is *Lie Machines: How to Save Democracy from Troll Armies, Deceitful Robots, Junk News Operations, and Political Operatives*. He blogs at www.philhoward.org and tweets from @pnhoward.

ABOUT THE OXFORD COMMISSION ON AI AND GOOD GOVERNANCE

The mission of the Oxford Commission on AI and Good Governance (OxCAIGG) is to investigate the artificial intelligence implementation challenges faced by governments around the world, identify best practices for evaluating and managing risks and benefits, and recommend strategies for taking full advantage of technical capacities while mitigating potential harms of AI-enabled public policy. Drawing from input from experts across a wide range of geographic regions and areas of expertise, including stakeholders from government, industry, technical and civil society, OxCAIGG will bring forward applicable and relevant recommendations for the use of AI for good governance.

oxcaigg.oii.ox.ac.uk