DATE: MARCH 18, 2019
TO: AUSTRALIAN HUMAN RIGHTS COMMISSION AND
THE WORLD ECONOMIC FORUM
FROM: THE MONTREAL AI ETHICS INSTITUTE

AUTHORS:
ABHISHEK GUPTA, FOUNDER
MIRKA SNYDER CARON, ASSOCIATE

Re: Artificial Intelligence: Governance and Leadership White Paper 2019 - Response to Consultation process

Executive Summary:

The following paragraphs summarize prioritized comments from the Montreal AI Ethics Institute's ("MAIEI") pertaining to the Australian Human Rights Commission White Paper.

If a central organization is to be established to play the role of promoting responsible innovation in AI and related technologies (the "Responsible Innovation Organization" or "RIO"), it will be very important for this organization to have public consultations be an essential part of its policy making. From our experience at the MAIEI, we have found this to be particularly effective in unearthing solutions that are interdisciplinary and contextually and culturally sensitive as well.

In the context of the RIO creating multi-stakeholder dialogue, it is the strong recommendation of the Montreal AI Ethics Institute that public consultation and engagement be a key component because it helps to surface interdisciplinary solutions, often leveraging first-hand, lived experiences that lead to more practical solutions. Additionally, such an engagement process at the grassroots level increase the degree of trust and acceptability on the part of the general public[14,22] since they would have played an integral part in the shaping of the technical and policy measures that will be used to govern the systems that they are going to be affected by.

Apart from setting up an RIO, it will be essential to ensure it be able collaborate with those existing organizations that we have listed below so as to not duplicate efforts or have to re-learn things that those organizations already have years of experience in. In fact, it would be great to have a system, whereby there is a distributed intelligence of "experts" across these organizations (akin to liaisons of the RIO) that work at each of

these organizations and are able to coordinate the work across the RIO and all the other organizations.

Furthermore, the scale of financial commitment required should be high to allow for meaningful work to happen and to be able to engage in the hard, long-term but ultimately impactful work of public engagement on this and building of public competence in building responsible AI systems.

When thinking about approaches, solutions, frameworks for public, private industries, care needs to be taken to make sure that the solutions are not generic and are tailored per industry, perhaps even split by sub-industries, because it is the recommendation of the institute, based on experience, that the more nuanced and specific the advice is, the more applicable, practical and integrable it is, ultimately increasing the efficacy of the work of the RIO. However, considering AI may have an impact on all industries, it is our recommendation, at time of evaluation and implementation, to combine specific concrete solutions curtailed to an industry with a holistic approach, since it is possible to gain multiple industries' consensus on key ethical priorities and fundamental human values. The holistic approach, supported by increased collaboration and shared expertise between regulators, while taking public and industry feedback into account, will prevent the risk of applying a siloed industry-specific approach.

Finally, standardization without an appropriate understanding on the part of the layperson (which is commonly non-existent) is very difficult if not impossible. In fact, it is potentially more harmful to have certifications in place that purport to guarantee some adherence to a higher quality of product while preserving the rights of users, but are in effect only a hollow affirmation. For example, the Statement of Applicability[23], which is usually only revealed under an NDA, showcases the extent to which the standards were applied and to what parts of the system. So for example in

cybersecurity ISO 27001 when looking at whether a system is compliant or not, one can have the certification but that doesn't mean that all the components of the system are covered in the evaluation to obtain that certification. In fact that SoA is what tells you which parts of the system were evaluated to grant the certification.

1) What should be the main goals of government regulation in the area of artificial intelligence?

We identify 5 main goals to government regulation for artificial intelligence:

1. Identify non-exclusive prohibited and unacceptable market or industry behaviours.

2. Clarify expectations relative to authorized and acceptable behaviours' minimal standards or requirements towards consumers and citizens.

3. Ensure adequate enforcement and monitoring powers to a democratically delegated, independent administrative or regulatory authority including through implementing sufficiently deterring administrative penalties and civil liability schemes.

4. Ensuring that AI-enabled solutions uphold applicable human rights and legal protections for individuals, especially protecting those that are marginalized.

5. Offer a framework and regulations for redressal when individuals face unjust decisions from automated systems.

2) Considering how artificial intelligence is currently regulated and influenced in Australia:

(a) What existing bodies play an important role in this area?
(b) What are the gaps in the current regulatory system?

The following organizations[6] are aware of AI opportunities, risks and developments:

- Australian Cyber Security Centre
- Minister for Industry, Technology and Science
- Australian Human Rights Commission
- Department of Education and Training
- Bureau of Communications and Arts Research
- Office of the Australian Information Commission- Privacy Act 1988
- Australian Securities and Investments Commission
- Australian Prudential Regulatory Authority
- Australian Competition and Consumer Commission
- Ministerial Council on Consumer affairs, and other consumer protection agencies and organisations
- IP Australia
- CSIRO, Australia
- ANU's 3AI Program

However, considering AI systems applications are not industry-specific and require multidisciplinary expertise and feedback from all population strata,

it is our recommendation to ensure added collaborative networking between existing regulators in any industry through the implementation of an independent and autonomous public entity. For example, it could become relevant for the Clean Energy Regulator or other ministry of Energy, to become involved in AI impact assessments, considering the potential increase in energy footprint. AI applications require massive compute for training and have large data storage requirements, and both may have large energy footprints. As such, consideration should be given to ensure the sustainable development of this technology.

We refer you to the mission and objective of the Global Financial Innovation Network[24] and recommend an analogous application at first the local, regional and national levels, and then on the international level[9,10,11]. Such collaboration should be predicated on sharing learnings and derived best practices in a timely manner and should be facilitated by experts that straddle technical and social sciences, especially when addressing the rapid pace of research and development in AI.

3) Sectors with gaps that need to be bridged[1-16]:

1) Healthcare: More dialogue is needed between patients and clinicians about AI systems design and use. There is a need to "bridge the translational gap"[15] to avoid patient confusion. New regulation should encourage engagement in the designing of AI systems, as well as to provide adequate disclosure of AI systems functions, limits and objectives, and reasonable expectations as to outcomes in a clear, simple and understandable language for patients concerned.

While anonymization and pseudonymization are usually the techniques that are deployed to protect data, there have been many instances where

such techniques have shown to be broken, see examples of AOL[25], Netflix[26] and Strava[27]. Instead we recommend the use of differential privacy[8] which offers robust mathematical solutions on the privacy of individual's data, even in the face of potentially unlimited access to auxiliary data sources to leverage the mosaic effect. Regulation should encourage "privacy and security by design" offering built-in privacy settings in the systems at time of design and prior to scaling. Opt-in consent of patient for algorithmic use of patient records and health profile ought to be implemented, including easy, real-time opt-out mechanisms at patient's will, to ensure adequate empowerment. If patient requires opt-out and continuous access to datasets is required to avoid damaging algorithms' accuracy, regulation should require such data to be appropriately anonymized with nil trace-back possibility at time of patient request.

2) Banking[18]: It is our general understanding that the government is intending to open access to banks' datasets which have been guarded by financial institutions, to level the playing-field in terms of market power with new emerging players and other constituents and digital services' providers, somewhat akin to the PSD2 "open banking" strategy of the European Union. Considering the traditionally sensitive and confidential nature of financial personal information, consideration should be given to ensure data is shared with trustworthy and cybersecurity-adhering data providers for AI systems, and ensure both legal and technical requirements to ensure appropriate enlightened consent is provided by the individual concerned, and that the individual may opt-out from consent at any time without significant negative impact on products and services provided.

3) Civil Liability and Accountability: Although the Australian Government state's existing legislation is technologically-neutral and provides sufficient protection and recourse to Australians[4], some specific amendments should be considered to better help frame common law claims in

emerging technologies such as AI. There is an existing ambiguity pertaining to civil liability and accountability allocation when an algorithm provides outcomes detrimental to a specific individual or to a group of people, such as vulnerable persons or minorities, in particular as to whether to treat it as a defective product or not, and whom to designate as responsible for the damages suffered throughout the chain of relevant stakeholders, such as the designer, programmer, seller, purchaser, or even the user of the AI system, and to what extent. Further considerations as to lifting of the "corporate veil" to target personal civil liability in cases of gross negligence, malicious intent or recklessness ought be clarified, implemented or excluded explicitly in AI systems scenarios.

Depending on the auditability of a given AI system, the "black box" problem[28] may increase causality link difficulties in terms of providing sufficient evidence on a balance of probabilities. As such, consideration should be given whether presumption benefitting the concerned damaged individual be specified, to relieve the burden of proof. Otherwise, clear indication whether or not a strict liability regime or a fault-based regime is to be expected should be considered.

Consideration should be given to monetary publicly-funded or privately-implemented compensation schemes for individuals affected by illegal discrimination or other negligent or defective algorithmic outcome, as well as possible insurance schemes.

4) Public surveillance of Australians: Considering recent developments facial-id, voice-id, and other biometrics-based and smart-home AI systems, authorized and prohibited uses of such systems ought to be clarified through amended regulation with a non-exhaustive list of examples, when used by either police and other public government entities, as well as private organisation for business uses. Considerations ought to be given to

include both civil and criminal penalties in case of unauthorized uses, as well as ensuring appropriate education, disclosure, awareness and empowerment of a consumer or citizen when using a product which could technically be used for surveillance purposes.

5) Public administrative decisions affecting Australians and private industry entities: As automation enters decision-making processes by public or regulatory entities which may have significant effect on an individual or private organisation, serious consideration ought to be given to regulation amendments or guidelines to ensure fair trial and hearing (*audi alteram partem*), as well as commit to right to object or to contest the algorithmic decision, right to be heard, and right to appeal the decision to a human bench. As an example, we refer you to the Directive on Automated Decision-Making of Canada[12], which also ensures for example peer-reviewed AI impact assessments depending on the level of risk.

Where legal predictive analytics might be implemented within the judicial system, even more stringent requirements of disclosure and adequate legal and technical safeguards to such judicial rights should be regulated and implemented.

6) AI Testing: AI systems are certainly being tested on present targeted sample groups of Australians consumers possibly in stealth mode, before being scaled systemically. Considerations should be given to regulate the manner in which AI systems may be tested, for instance at the time of "proof-of-concept" stage; prior on differentially private datasets; on clearly identified and consenting small group of existing employees or customers; with various measures and safeguards in place prior to systemic scaling.

7) AI Identification: Various jurisdictions, such as California[29], are working on bills to ensure a prospective customer is made aware when conversing or dealing with AI systems such as chatbots or other automated entities, instead of a human. Consideration should be given to ensure appropriate disclosure of the nature of the agent with whom an individual is communicating or transacting, as well as about objective, limits, possible outcomes and risks.

8) Manipulation of people through social media: There is ample evidence that people can be nudged via social media to alter their voting patterns[30], thus threatening the fundamental tenets of a well-functioning democracy. Aside from that, hyper-personalization of content using machine learning techniques can nudge purchasing patterns, interactions with other individuals, what kind of news is consumed, etc. ultimately having a very deep impact on the existence of an individual. Keeping this in mind, we recommend that specific policies and guidelines be made to address how these platforms operate, how they utilize data, how they target consumers, and how they present their terms and conditions. We also recommend that there be development of techniques and metrics that can help measure the impact of how much nudging happens to individuals and if that falls within appropriate limits. Furthermore, providing added regulatory clarity pertaining to the expected balance between freedom of expression and censorship, and pertaining to mitigating the social contagion and sharing of fake news, by bots on social media with both legal and technical measures should be considered.

9) Intellectual Property[5-19]: Added clarity as to applicable intellectual property protection pertaining to AI could be considered, to mitigate unnecessary litigation, invalidations or to augment leading innovation in

Australia. For instance, in certain jurisdictions, mathematical formulas or algorithms per say are very difficult and perhaps impossible to protect under patent or copyright law, unless combined with new and innovative elements. In the EU, it has been made clear that algorithms cannot be protected under IP law per say. Harmonization measures with WIPO should also be considered.

4) Would there be significant economic and/or social value for Australia in establishing a Responsible Innovation Organisation?

Significant GDP growth is estimated in coming years globally for jurisdictions which have leading technical, ethical and regulatory expertise in managing AI systems. We can look to the European Union where with the introduction of GDPR[31] and partly driven by a larger public awareness of the importance of privacy and digital rights, users are moving towards products and services that offer stronger security and privacy measures. We believe that there will be a similar wave for AI solutions when it comes to ethical, safe and inclusive design, development and deployments. Both private and public investments are being made in the technology. Enabling an RIO to increase such expertise, to increase collaboration between industry-expert national regulators, to monitor the degree of penetration of AI systems within society as well as to provide periodic reports on AI impact and adaptability of workforce to the government, to prevent misuse and to enforce compliance on negligent, malicious or reckless entities or individuals, should bring both economic and social value to Australia. Some other jurisdictions have already developed

committees, ministries and other public entities to this effect (eg. USA, UK, and Canada).

5) Under what circumstances would a Responsible Innovation Organisation add value to your organisation directly?

The Montreal AI Ethics Institute is a registered non-profit organisation which has the mission of defining humanity's place in world increasingly driven and characterized by algorithms through concrete and tangible solutions. As such, we monitor technical, ethical and regulatory developments in countries around the world to provide expert guidance to public entities and to increase awareness and education of the public through free public competence-building sessions, workshops and published articles. Our strict focus is on applied, practical and tangible solutions, both technical and social, that meaningfully address the issues when it comes to the societal impacts of AI[32]. Our work has been recognized by governments from North America, Europe, Asia and Oceania.

Were Australia to designate an RIO, we would benefit greatly in identifying centralized government expertise directly relevant to our mission, enabling us and other organisations to optimize information gathering on such developments in Australia as well as enable better communication and collaboration with such RIO for future work. Additionally, we believe that a close collaboration between the Montreal AI Ethics Institute and the RIO in Australia would enable our organization to test and validate hypotheses, run applied experiments and work with the RIO to share

lessons learned, experience and expertise that the institute has gained through work with other public entities across the world.

6) How should the business case for a Responsible Innovation Organisation be measured?

Key performance indicators or metrics could be identified, such as, without being limited to (inspired by the Canada Privacy Commissioner Annual report to Parliament structure):

complaints accepted/cases
closed cases through early resolution
closed cases through standard investigation
breach reports
AI Impact Assessments received
Advice provided to public sector organizations (review or consultation)
Advice provided to private sector organizations (review or

consultation)
Bills and legislation reviewed for AI implication
Appearances with parliamentary committees on AI matters for public and private sectors
Formal briefs submitted to Parliament on AI matters for public and private sectors
Speeches and presentations
public competence-building sessions run
people engaged in public competence-building sessions
Visits to website
Blog visits
Social media messages sent (eg. Twitter, Facebook, etc.) and #followers
Publications distributed

News releases and announcements
Newsletter subscriptions
International regulators meetings and public endorsements of consensual declaration for AI ethics (eg. See “Declaration on Ethics and Data Protection in Artificial Intelligence”[33])

7) If Australia had a Responsible Innovation Organisation:

(a) What should be its overarching vision and core aims?

(*The following are inspired from the Global Financial Innovation Network consultation paper[34], feedback and next steps).

Vision: We suggest that the vision of the RIO should be to ensure fair, ethical and responsible design and use of AI systems within Australia for the best interests of Australian citizens, and other international stakeholders in the context of Australian-based crossborder solutions.

Core aims: We suggest three core aims of the RIO:

- To act as a catalyser for increased networking and collaboration amongst regulators of different industries in Australia which are

presently or are going to be increasingly impacted by AI developments, by sharing the experience of innovation in different markets, and to provide publicly accessible regulatory information for both private and public entities.

- To provide a forum for joint innovation work and collaborative knowledge sharing and lessons learned between national and international regulators, but also with relevant private entities developing or using the technology.

- To provide voluntary firms with a collaborative environment in which to trial national and crossborder solutions (eg. Robotland (South Korea), sandbox, access to differentially private databases for training and testing for pilot or “proof-of-concept” phase, etc., templates for AI Impact Assessments and Human Rights Due Diligence pipelines or checklists, etc.)

(b) What powers and functions should it have?

We identify multiple powers and functions the RIO should be bestowed with to ensure optimal efficiency:

- Review, verification and recommendation functions: if a confidential self-disclosure scheme for private organisation on the level of compliance with AI legislation or standards were to be implemented, the RIO should review and verify statement of compliance and recommend technical or operation actions or measures to be followed as required. (*Mandatory periodic disclosure by market participants would enable better statistical representation of AI

developments). Other examples of recommendations for technical and policy work could include “Datasheets for Datasets”[20], or “Model cards for model reporting”[21].

- Monitoring powers: the RIO should have the discretionary powers to request access to AI systems details or processes when it receives a complaint or when it has reasonable doubt or suspicion that a breach has occurred or that risk of breach is imminent;

- Enforcement powers: Were a specific legislation be put in place pertaining to responsible design and use of AI, or were duplicated civil and penal proceedings and forum shopping risks be sufficiently mitigated by appropriate measures and safeguards for the RIO to have overarching nationwide powers, the RIO should be provided with clear enforcement powers to ensure compliance to the responsible standards or regulatory requirements, through an administrative decision-making procedure combined with monetary administrative penalties, injunctive relief and powers to prohibit access to the market or to have activities ceased by concerned private organisation or individual.

- Publication function to the public: the RIO should provide ongoing information to the public about its activities and expertise.

- Publication function to the parliament and government: the RIO should provide ongoing urgent information, or at least one annual report pertaining to the status of AI developments, responsible innovation in this field and degree of societal penetration of these systems.

- Networking, expertise, knowledge sharing and collaboration functions: the RIO should ensure regulators of all industries in Australia are provided with information about AI systems and developments to ensure adequate expertise across the country. Additionally, the RIO should support specific research and development of tools, techniques and frameworks, both technical and non-technical for the responsible development of AI, covering areas like bias, transparency, explicability, accountability, interpretability, etc. As another example of work towards the public benefit, the subjects of AI systems should have the option to receive advice/actions that they can take to "improve their conditions" as judged by the system, i.e. the subjects should be able to, through their actions, work their way out of algorithmic determinism[35]. It should also be accessible to laypersons so that the redressal mechanisms are inclusive.

- Power to share private organisation information with other national and international regulators on a "necessary" basis: Considering the potential crossborder application of AI systems and the high degree of interconnectedness of the global economy to date, the RIO should be provided with the power to share information on a given complaint or private organisation to other regulators which would otherwise be held confidential when it is necessary to prevent or mitigate systemic negative effect of an AI system in other jurisdictions.

- Personal judicial immunity but for gross negligence or corruption: The RIO members or employees should be able to do their job

without fear of becoming personally sued in court, as is standard, but for evidence that they proceeded in their tasks with gross negligence or show evidence of corruption.

(c) How should it be structured?

The RIO should be independent and autonomous, but should report to Parliament and Government about its activities, budget plans and annual strategy, but for delegated administrative powers amended through legislation.

The RIO should ensure administrative decision-makers remain independent, unbiased and impartial in their role. As such, standard measures and safeguards to ensure absence of conflicts of interests and avoidance of political influence should be put in place.

The RIO should be composed of both permanent and temporary representatives. By permanent we first mean elected administrative representatives and decision-makers, as well as regulatory and technical experts for a mandate of 4 to 5-year, renewable once. By temporary we mean designated representatives from other regulatory entities to share knowledge openly. The RIO should also consist of members from the public-at-large that would like to serve on the regulatory and technical committees within the RIO. Leveraging grassroots expertise will not only serve the function of being more inclusive but will also encourage the development of public competence and public engagement will increase the trust and acceptability of solutions from the RIO.

(d) What internal and external expertise should it have at its disposal?

Subject to budget constraints, optimal scenario would be for the RIO to have internal regulatory, ethical and technical expertise pertaining to AI systems.

If this is not possible, then the RIO should put in place public contracts agreements to obtain this expertise through outsourcing with academic institutions, ethics institutes, legal firms and other robotics and AI standard-setting organisations, to guide them in their policy recommendations and decision-making processes.

(e) How should it interact with other bodies with similar responsibilities?

Care should be taken to ensure there are no duplicated legal or administrative decision-making processes or regulatory legislation-based mandates, as this will only increase costs and may create confusion for market participants and public were regulators to provide different responses to the underlying national responsible AI strategy. Furthermore, this may increase activities of forum shopping if a regulator or other body appear more favourable than another.

The RIO should be as transparent about its activities, compliance expectations, decision-making processes and expertise as possible, with the public and other bodies, and make as much information freely and publicly available as possible, subject to confidential information about private organisations under scrutiny or investigation in specific cases, as the case may be. Freely accessible information and active knowledge sharing would meet its core functions as identified above.

(f) How should its activities be resourced? Would it be jointly funded by government and industry? How would its independence be secured?

On a national level, funding and resources should first be provided through specific allocation from the national annual budget.

Otherwise, second alternative funding could be considered through a formal percentage-based or fixed-amount based contributions from private organisations designing or developing AI to be strictly used. A trust fund could be put in place to ensure funds are used solely for specific objectives. It would also be possible to re-allocate a portion of these private sector contributions to a compensation scheme, to mitigate damages to Australian consumers/citizens in case an AI system has a detrimental systemic national effect. The important point to note here will be that funding from private organizations not influence the work of the RIO, especially if those organizations are under scrutiny.

Finally, additional funding could perhaps be found with international non-governmental organisations looking to allocate funds to ensure adequate regulatory and ethical protection in specific jurisdictions, or to support ongoing regulatory projects to frame responsible AI innovation.

(g) How should it be evaluated and monitored? How should it report its activities?

See suggestions under sections 5, publication functions under 6b), and 6c).

Authors:

Abhishek Gupta: Founder of Montreal AI Ethics Institute and a Machine Learning Engineer at Microsoft where he serves on the CSE AI Ethics Review Board. His research focuses on applied technical and policy methods to address ethical, safety and inclusivity concerns in using AI in different domains. He has built the largest community driven, public consultation group on AI Ethics in the world that has made significant contributions to the Montreal Declaration for Responsible AI, the G7 AI Summit and the European Commission Trustworthy AI Guidelines. His work on public competence building in AI Ethics has been recognized by governments from North America, Europe and Asia.

Email: abhishek@montrealethics.ai

Mirka Snyder Caron: Associate at the Montreal AI Ethics Institute, as well as certified Quebec and Ontario lawyer who worked in-house at National Bank of Canada for 4 years addressing issues like IP, Fintech and Big Data. Her thesis is on the "Transformative Effect of AI on the Banking Industry: A Legal and Economic Analysis", to be published in April 2019 in the Canadian Banking Law and Finance Review.

Email: mirka@montrealethics.ai

Montreal AI Ethics Institute: Our mission is to help define humanity's place in a world increasingly characterized and driven by algorithms. We do this by creating tangible and applied technical and policy research in the ethical, safe and inclusive development of AI. Our unique advantage is that Montreal is situated at the leading edge of technical research while leveraging strong Canadian values of diversity and inclusion.

Website: https://montrealethics.ai

References:

1- ANDERSEN, Lindsey, “Human Rights in the Age of Artificial Intelligence”, Accessnow, 2018, https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf (visited on March 17, 2019)

2- AI Now Institute, https://ainowinstitute.org (visited on March 17, 2019)

3- AI Now Institute, “AI Now Report 2018”, https://ainowinstitute.org/AI_Now_2018_Report.pdf (visited on March 17, 2019)

4- Australian Government, “Australia’s Tech Future: Delivering a Strong, Safe and Inclusive Digital Economy”, 2018, 52 pages.

5- Australian Government, IP Australia, “IP Australia and the Future of Intellectual Property”, 2017, https://www.ipaustralia.gov.au/sites/g/files/net856/f/ip_australia_and_the_future_of_intellectual_property.pdf (visited on March 17, 2019)

6- Australian Securities and Investments Commission, “Other Regulators and Organizations”, https://asic.gov.au/about-asic/what-we-do/our-role/other-regulators-and-organisations/ (visited on March 17, 2019)

7- CHESSEN, Matt, “The AI Policy Landscape”, Medium, 2017, https://medium.com/artificial-intelligence-policy-laws-and-ethics/the-ai-landscape-ea8a8b3c3d5d (visited on March 17, 2019)

8- ELAMURUGAIYAN, Aaruran, “A Brief Introduction to Differential Privacy”, Georgian Partners, 2018, https://medium.com/georgian-impact-blog/a-brief-introduction-to-differential-privacy-eacf8722283b (visited on March 17, 2019)

9- Financial Conduct Authority, “Global Financial Innovation Network Consultation”, 2018, https://www.fca.org.uk/publications/consultation-papers/global-financial-innovation-network (visited on March 17, 2019)

10- Financial Conduct Authority, “Terms of Reference for Membership and Governance of the Global Financial Innovation Network (GFIN)”, 2018, https://www.fca.org.uk/publication/mou/gfin-terms-of-reference.pdf (visited on March 17, 2019)

11- Global Financial Innovation Network (GFIN), “Consultation Paper”, 2018, https://www.fca.org.uk/publication/consultation/gfin-consultation-document.pdf (visited on March 17, 2019)

12- Government of Canada, “Directive on Automated Decision-Making” http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592 (visited on March 17, 2019)

13- GUPTA, Abhishek, “Legal and Ethical Implications of Data Accessibility for Public Welfare and AI Research Advancement”, 2018, https://towardsdatascience.com/legal-and-ethical-implications-of-data-accessibility-for-public-welfare-and-ai-research-advancement-9fbc0e75ea26 (visited on March 17, 2019)

14- GUPTA, Abhishek, “Inclusive Design - Methods to Ensure a High Degree of Participation in AI Systems”, Connected Life 2018: Information Control, Oxford Internet Institute, 2018. (visited on March 17, 2019)

15-Healthcare IT, “Bridging the Human-Machine Gap in AI-Augmented Healthcare”, 2018, https://www.healthcareit.com.au/article/bridging-human-machine-gap-ai-augmented-healthcare (visited on March 17, 2019)

16- MERCHANT, Brian, “Top 5 Most Worrying Trends in Artificial Intelligence”, 2018, https://www.gizmodo.com.au/2018/12/the-five-most-worrying-trends-in-artificial-intelligence-right-now/ (visited on March 17, 2019)

17- ROSENBLATT, Gideon, GUPTA, Abhishek, “Artificial Intelligence as a Force for Good”, Stanford Social Innovation Review, 2018, https://ssir.org/articles/entry/artificial_intelligence_as_a_force_for_good# (visited on March 17, 2019)

18- SNYDER CARON, Mirka, “The Transformative Effect of AI on the Banking Industry: a Legal and Economic Analysis”, Banking Law and Finance Review, 2019, (to be published, April 2019), 46 pages.

19- JEFFRIES, Andrea W., TAIT, Emily J., “Protecting Artificial Intelligence IP: Patents, Trade Secrets, or Copyrights?”, 2018, https://www.jonesday.com/protecting-artificial-intelligence-ip-patents-trade-secrets-or-copyrights-01-09-2018/ (visited on March 17, 2019)

20- GEBRU, Timnit, MORGENSTERN, Jamie, “Datasheets for Datasets”, 2018, https://arxiv.org/abs/1803.09010 (visited on March 17, 2019)

21- MITCHELL, Margaret, WU, Simone, “Model Cards for Model Reporting”, 2019, https://arxiv.org/abs/1810.03993 (visited on March 17, 2019)

22- GUPTA, Abhishek, “AI Ethics: Inclusivity in Smart Cities”, Montreal AI Ethics Institute, 2018, https://medium.com/montreal-ai-ethics-institute/ai-ethics-inclusivity-in-smart-cities-6b8faebf7ce3 (visited on March 17, 2019)

23- KOSUTIC, Dejan, “The importance of Statement of Applicability for ISO 27001”, Advisera, 2017, https://advisera.com/27001academy/knowledgebase/the-importance-of-statement-of-applicability-for-iso-27001/ (visited on March 17, 2019)

24- Global Financial Innovation Network, https://www.fca.org.uk/firms/global-financial-innovation-network (visited on March 17 2019)

25- BARBARO, Michael and ZELLER JR., Tom, “A Face Is Exposed for AOL Searcher No. 4417749”, New York Times, 2006, https://www.nytimes.com/2006/08/09/technology/09aol.html (visited on March 17 2019)

26- NARAYANAN, Arvind and SHMATIKOV, Vitaly, “Robust De-anonymization of Large Sparse Datasets”, University of Texas at Austin, 2007, https://www.cs.cornell.edu/~shmat/shmat_oak08netflix.pdf (visited on March 17 2019)

27- HERN, Alex, “Fitness tracking app Strava gives away location of secret US army bases”, The Guardian, 2018, https://www.theguardian.com/world/2018/jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases (visited on March 17 2019)

28- BATHAEE, Yavar, “The Artificial Intelligence Black Box and the Failure of Intent and Causation”, Harvard Journal of Law and Technology, 2018, https://jolt.law.harvard.edu/assets/articlePDFs/v31/The-Artificial-Intelligence-Black-Box-and-the-Failure-of-Intent-and-Causation-Yavar-Bathaee.pdf

29- GERSHGORN, Dave, “A California law now means chatbots have to disclose they’re not human”, Quartz, 2018, https://qz.com/1409350/a-new-law-means-californias-bots-have-to-disclose-theyre-not-human/

30- BISWAS, Aindrila, INGLE, Nikhil and ROY, Mousumi, “Influence of Social Media on Voting Behavior”, Journal of Power, Politics & Governance, 2014, http://jppgnet.com/journals/jppg/Vol_2_No_2_June_2014/7.pdf

31- PARDES, Arielle, “What is GDPR and why you should care?”, Wired, 2018, https://www.wired.com/story/how-gdpr-affects-you/

32- Montreal AI Ethics Institute, About, https://montrealethics.ai/about/

33- “Declaration on ethics and data protection in Artificial Intelligence”, International Conference of Data Protection and Privacy Commissioners, 2018, https://icdppc.org/wp-content/uploads/2018/10/20180922_ICDPPC-40th_AI-Declaration_ADOPTED.pdf

34- Global Financial Innovation Network Consultation Paper, 2018, https://www.fca.org.uk/publication/consultation/gfin-consultation-document.pdf

35- POLONSKI, Vyacheslav, “Algorithmic determinism and the limits of artificial intelligence”, Oxford Internet Institute, 2016, https://www.oii.ox.ac.uk/blog/algorithmic-determinism-and-the-limits-of-artificial-intelligence/