

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,600

Open access books available

138,000

International authors and editors

175M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Introductory Chapter: AI's Very Unlevel Playing Field

Ali Hessami and Patricia Shaw

1. Introduction

There are many great initiatives happening in the space of AI and data ethics. There are a variety of high-level principles, process and procedural standards, risk and impact assessments, certification, and audit all in the making. However, to move from voluntary adoption and inconsistent application of such great works, will require robust policy and ultimately law.

Calling for the need for the regulation of AI, is not innovation stifling. Instead it has the potential to birth an industry, create a level of reliability and safety for people and planet that has not been previously secured, to embed human dignity, human flourishing, human autonomy, freedom of choice and non-discrimination into AI design, development, deployment, monitoring and decommissioning. AI is formed in a lifecycle but implemented and operationalised in a diverse ecosystem which contains contrasting and competing interests, where context truly matters. Therefore, no regulatory response should avoid this complexity but tackle the challenge head on in a manner which enables it to flex to a constantly changing technological environments and be sufficiently agile and adaptable to be future-proof. It will require learning lessons from the history of AI products and services, and a deep dive into the possibilities of existing and emerging technologies that lie before us.

We need to understand both the risk and the likelihood of the risk impact occurring in the short, medium and long term, and how this risk and its impact changes from context to context, country to country, culture to culture. Diversity, equity and inclusion matter.

We must evaluate our infrastructure, our governance frameworks, and design a regulatory response that is fit for purpose both from the top down and bottom up. Better business, better outcomes, better society, can all be born out of greater stakeholder engagement and participatory governance in AI. Recurrent and dynamic feedback should be our weapon of choice to head not just legal but also ethical risk off at the pass, to abate biased or unfair and exclusionary outcomes, technologically disguised anti-competitive behaviours, unintentional consumer and citizen harm, indirect and inadvertent discrimination, and ultimately unconscious human rights impacts and infringements.

We cannot preserve the status quo otherwise we will simply sleepwalk into making the same mistakes of the past, embedding historic and systemic attributes and risks. Compliance on a mere voluntary or “soft law” basis will not simply cut the mustard. Neither will a siloed jurisdictional approach to AI regulation. Cohesion, cooperation, collaboration will be key for any new regulatory system which seeks to transcend regulatory arenas and cross national borders.

The development of a globalised AI ecosystem sets the requirement for an umbrella international regulatory response. The global–local dichotomy and paradox

can no longer be ignored. Utilising generic data and AI tools to apply generalisations from one jurisdiction to the next must be called out for what they are “irresponsible AI” based on a lucky guessbet. This is where the importance of localism, culture and context will come into their own. An international regulatory response will be most effective where it brings parties together with a consistent nomenclature with regularisation tools such as standards, certification and audit which must be built with globally diverse and inclusive actors. It will not be at its best by effectuating what can only be described as AI ethics colonialism, seeking to apply one set of ethics to all contexts. This is why application and enforcement should be (and ought always be) best left in the hands of the relevant jurisdiction(s). This way AI ethics can be both contextually. Culturally and equitably applied.

2. The case for ethics

“[There is a] need for Governments, the private sector, international organizations, civil society, the technical and academic communities and all relevant stakeholders to be cognizant of the impact, opportunities and challenges of rapid technological change on the promotion and protection of human rights, as well as of its potential to facilitate efforts, to accelerate human progress and to promote and protect human rights and fundamental freedoms” [1].

Ethics is a forerunner to legislation. It is the ethical dilemmas that we face as a society that prompt the need for new law, where existing laws do not or cannot fill the gaps. Law provides certainty and resolve to the social problem faced. At some point in every generation, a society has to decide what is acceptable or not acceptable behaviour or outcomes. This generation no less than any generation that has gone before it. Simply this time it concerns AI, or more pertinently the use of big data and algorithmic intelligent and (semi-) autonomous systems, the control (or not) the designers, developers and deployers, and those that monitor their performance, have over and concerning the outcomes.

The AI ecosystem and supply chain are complex which makes legislating it and ensuring that it is futureproof too so very tricky.

AI ethics itself has so far proven popular. It has certainly raised the issue of “trustworthy” [2]. AI not only at a national but international scale. The challenge is, has AI ethics alone really changed anything at all! Being a leader and a responsible AI advocate can be a real competitive advantage, but this is where operationalising AI ethics moves the goalposts from Advocate to Actor to Ambassador, where AI ethics and the governance that operationalises it can become a real innovation enabler.

Change must occur. But it can only do so if governments and businesses are willing to take the first steps to learn how to operationalise AI ethics, and embed agile and dynamic governance which works in harmony with its stakeholders.

AI ethics is not merely about securing privacy (or more pertinently data privacy) for end users. Although that is a step in the right direction. It is about creating an equitable digital society whereby human and organisational, socio- and technical tools work towards trustworthy human oversight, informed human agency, and good exercise of human autonomy. A move away from bias, underrepresented people groups and lack of diversity, towards fairer and non-discriminatory outcomes; allowing for appropriate process and procedural transparency as well as decision transparency, not just transparency of data, models and code, to ensure the necessary safeguards are in place to provide qualitative and quantitative assurance of safety and reliability, societal and environmental wellbeing; and (last but not least) knowing who, how, why and when someone should be accountable.

If this is to be operationalised at a national scale, government departments and businesses need to be given permission (and a good nudge) to allocate resource, time, effort and budget to AI ethics, its risk management and impact assessment, its governance, and ultimately its compliance. It needs elapsing of time and experience to move from competence, capability, capacity building to maturity.

Ethics alone lacks “teeth” and it is the obligatory requirement and the enforceability that the law offers that make ethics translation into law so attractive. It moves voluntary codes of conduct and ethical principles to a stable and more sure footing when it is mandated and enforceable through legislation or regulation.

This needs a suitable national regulatory environment, that recognises how and where AI's impact interplays with existing law, and how legal and regulatory gaps can be plugged.

If this is to be operationalised at an international scale, AI ethics will need a common language and to be decolonialised. There is no one size fits all approach to this global ethical dilemma. This is a global–local problem and needs international cooperation and collaboration, but grassroots understanding of the problems it presents and the people it impacts in a given jurisdiction, sector or cultural space. The impact on the planet is a problem for us all, so making AI sustainable and handling the issue holistically so that we do not perpetuate existing environmental discrepancies and mismanagement through geo-political division.

International bodies and national governments being open, and regulators and regulated businesses being responsive will be key as we move from the age of AI discovery into the age of AI implementation.

The UN Human Rights Council's report [3], “The right to privacy in the digital age”, attempts at identifying and clarifying principles, norms and best practices relating to the promotion and protection of privacy rights in the digital age that also addresses the responsibilities of businesses enterprises in this context. The report provides guidance on how to address the emerging and pressing challenges to the privacy rights in a pervasively digital world. It explores the trends and concerns that interfere with privacy from a growing digital footprint to state surveillance and describes the responsibilities of the states to recognise, respect and protect the citizens' rights to privacy and the necessity for oversight and safeguards. The report also defines responsibilities for the business enterprises including respect and observance of human rights and the underpinning policies and procedures appropriate to the context, size and nature of its operations including due diligence in identifying and addressing the impact of their operations on human rights. The UN High Commissioner for Human Rights makes a number of recommendations aimed at the states and the business enterprises in recognising, evaluating and addressing the full implications of new data driven/intensive technologies on the human rights of the citizens.

On the business front, there's a rather unexpected upturn trend in the environmentally sustainable and ethical index funds outperforming the traditional investment funds even considering the impact of the pandemic on the markets [4]. The ethical index funds are now regarded as mainstream, a position that has traditionally been regarded as niche and at best minority.

The ethical index funds launched by Vanguard [5], one of the biggest global fund managers are offered under the Environmental, Social, Governance (ESG) class of funds that are branded as aligned with investor ethical principles. These are the three categories of ethical criteria that sets ESG funds apart from the traditional high return regular index sectors and track specific stock market indices that exclude companies that do not meet the independently established ESG norms and standards. Similar trends are emerging in the climate focused funds.

Morningstar, the global research agency that examined 745 sustainable funds when compared with 4,150 traditional funds surprisingly found that the majority of ethical/sustainable funds matched or outperformed returns on traditional funds in the UK or abroad over multiple time horizons [4]. This situation continued even during the COVID crisis. Another interesting facet is the longevity of the sustainable funds, effectively doing better over longer periods without the quiet removal or merger with better performing funds that's practiced by fund managers to boost the overall performance figures.

Overall, funds with a better and robust environmental, social and governance (ESG) focus and strategic management, are seen by industry observers as responsible investment that are better performers financially thus having a positive bottom line impact whilst aligning with social and ethical values.

3. Need for a balanced approach

There is an increasing body of ex-ante AI risk and impact assessment lists and questionnaires and standards for internal quality control and best practice processes, and for ex-post AI audit.

In the teachings of the prophet Zoroaster (630–550 BC) the universe is portrayed as a battle ground for good and evil [6]. Taoism, that also emerged around the sixth century BC believe that ultimate reality is beyond the capacity of reasoning and rational thought and interpreted the changes in nature as a result of interplay between polar opposites of *yin* and *yang* implying a belief in the unity of opposites. In a similar analogy to the Zoroastrian forces of good and evil, the Taoists strive to attain and maintain a dynamic balance between the polar opposites of *yin* and *yang* which are seen as a spontaneous and innate tendency in all things.

The traditional approach to the identification, evaluation, and management of risks (potential losses arising from hazards) and rewards (gains and benefits arising from the exploitation of opportunities) is that of minimization and maximisation whereas these are essential attributes in any facet of life, as recognised and practiced by the ancient wisdom of Zoroastrianism and Taoism. A holistic and balanced approach to the understanding and rational impact assessment of Autonomous Decision Making and Algorithmic Learning Systems (ADM/ALS) is to treat the hazards and opportunities as intertwined and omnipresent albeit associated with inherent ontologic and epistemic uncertainties.

This holistic framework is shown in **Figure 1** where typically hazards and threats are transformed into a spectrum of potential risks and opportunities into rewards/gains respectively [7]. The outcome is the spectrum and scale of risks and rewards that on balance informs the stakeholders in their desired and preferred decisions.

This framework provides a holistic, rational and unambiguous view of the key influencing factors in the impact assessment of ADM/ALS avoiding isolated treatment and confusing *upside* and *downside* terminology often employed to inadequately convey the same concepts or intent.

Such discussions of a holistic framework have led to an increasing need to ensure (and provide assurance of) oversight, to enable multi-disciplinary scrutiny of AI, to challenging the asymmetries of power between those organisations deploying AI systems (whether they be public sector or private sector), and those individuals, legal persons, people groups and wider society impacted by AI system outcomes. This has led the UK's ADA Lovelace institute to undertake a landscape review of algorithmic assessment and AI audit tools [8].

The challenge is that whilst seeking to balance tensions and trade offs and find tools, methods and approaches that can be operationalised to gain transparency

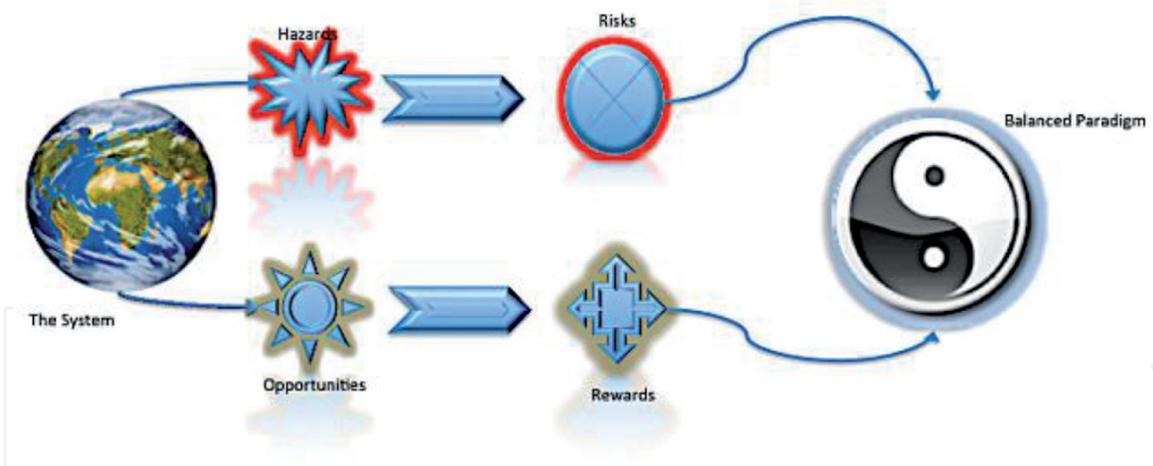


Figure 1.
A holistic risk–reward framework.

and hold AI systems (and the people behind them) to account, we need to find a common language, and an agreed taxonomy of terminology, that can not only cross language barriers but barriers of discipline too. There is still debate as regards how to define Artificial Intelligence itself. There is current debate as to whether ex-ante AI risk and impact assessments or ex-post AI audit should be looking at bias, unfairness or discrimination, or whether there is a case for all three areas to be within scope.

Ultimately the key is to ensure that no false positives or false negatives, no outliers or trends, no amount of data, no accuracy or inference, results in people being marginalised, excluded, rejected, or even expelled from operating and function in modern society. If people can or have been denied access to justice, social welfare, law enforcement, democratic engagement, employment, access to financial services, healthcare, education, or goods and services because of an AI system or as a symptom of all pervasive AI adoption without alternative or ability to opt out, we risk creating an ethical and societal divide. This requires algorithmic accountability, not least of all with the public sector [9].

In respect of the digital divide, according to the Office of National Statistics in the UK, in 2018 there were still some 5.3 million adults in the UK (amounting to 10.0% of the adult UK population) who were non-internet users. This simply provides a glimpse of the digital divide [10].

4. Pragmatic solutions

All tech solutionism aside, there is a place for human interventions, organisational approaches and socio-technical tools to develop and govern AI. There is no one size fits all approach. There is no one tool that can provide a silver bullet. It requires a holistic approach.

Understanding the purpose and outcomes to be achieved is a necessary first step. Many governments around the world are looking to algorithmic transparency to find ways of explaining automated decision making to its citizens. This on the one hand shows government to be open and accountable, but on the other hand it is a ruse to publicly legitimise their actions or inactions. Is not government responsible for the outcomes it creates in the public interest whilst also under a duty of care to ensure the safety of the wider public? If the public does not legitimise certain AI or ADM uses by government, what does that say to government in how it does or does not exercise its duty of care. How can we expect the government to fulfil its duty to

the masses without leaving the less represented and marginalised groups in society exactly that....marginalised!

Transparency in all its forms is an important key step but must be accompanied with meaningful stakeholder engagement. Transparency is the gateway to many of the other ethical principles, but for transparency to do its work, it must be explainable and understood in context in a way which is relevant to the recipients of the information – the message received is after all the message given.

Tools such as AI registers and risk analytics platforms are needed to accompany governance but more need to be done. In order for there to be a holistic and pragmatic approach, AI governance need to take into account human intervention, organisational processes as well as technological tools, especially those that increase our understanding and provide meaning and interpretation of what exactly goes on in that opaque box. This way ethics can be turned into something that is operational. It also has opportunity to legitimise governmental use of AI and to reaffirm their societal mandate to act in the public interests.

5. Current trends and way forward

The European Commission has made a brave and bold move to seek to regulate in the area of AI. In an effort to build an ecosystem of excellence and trust, it seeks to preserve European values and protect the Fundamental rights of European citizens. It's human centred approach to AI is to be applauded, especially as it seeks to provide a governance structure for AI, with scope for risk and impact assessment, adherence to standards and other voluntary codes of conduct, providing for conformity assessment (akin to product liability legislation) for those AI deployments which are deemed “high risk”.

Whilst this piece of legislation seeks to have extra-territorial effect like GDPR [X], it is not the GDPR of AI. Furthermore, it is a risk based, not principles-based piece of legislation like GDPR, but it does share something in common with GDPR: it is making the world's ears prick up. We may indeed see that all important “Brussels Effect” for AI governance crossing jurisdictional, geographical, and cultural divide, decolonising AI and AI ethics.

Barriers to global roll out and wider spread adoption of a regulatory approach such as this will be economic (determined by views of regulation stifling innovation), political (in the AI race), and will concern ethical disparities (public good versus equity and justice for the individual).

From a broader ethical perspective, three key areas of concern in development and deployment of ADM/ALS relate to Accountability, Transparency and freedom from unacceptable Algorithmic Bias. To this end, the IEEE-Standards Association has developed a suite of detailed criteria for evaluation, assessment and certification of these properties of ADM/ALS products and services under the “Ethics Certification Programme for Autonomous and Intelligent Systems” (ECPAIS). This programme [11] is a key facet of the IEEE-SA's Global Initiative and Ethically Aligned Design portfolio.

The three classes of ethical dysfunctions that may emerge in the embedding of ADM/ALS in products, systems and services require a systematic and credible independent evaluation and assurance to allay the public and private sectors' concerns and foster acceptance and deployment. To this end, IEEE-SA's suite of pragmatic and holistic certification criteria are now ready for deployment and tailoring for specific sectors and applications.

The high-level principles (Evaluation and Certification Factors) for each of the currently three ECPAIS suites are broadly defined as hierarchy of more detailed

factors and criteria (typically, 10–20 for each of the depicted high-level factors) which are S.M.A.R.T i.e. specific, measurable, achievable, realistic and timely at the pertinent system or component level.

Transparency relates to the criteria and values embedded in a system design and the openness and disclosure of choices and decisions made for development and operation. This applies to the entire ADM/ALS context of application for the product or service under consideration such as data sets and not restricted to technical and algorithmic aspects alone.

Accountability considerations concern the commitment by individuals and institutions involved in the design, development or deployment of ADM/ALS to remain responsible for the behaviour of the system as long as its integrity is respected. This is predicated on the recognition that the system/service autonomy and learning capacities are the result of algorithms and computational processes designed by humans and those humans should remain responsible for their outcomes. A key driver in accountability is explicit, sufficient and proper documentation and traceability for system design, development and deployment.

Algorithmic Bias relates to systematic errors and repeatable undesirable behaviours in an ADM/ALS that create unfair outcomes, such as granting privileges to one group of users over others where they are expected to be neutral and unbiased. This can emerge due to many factors, from the design of the algorithm influenced by pre-existing cultural or institutional practices, the decisions relating to the way data is classified, collected, selected or used to train the algorithm, the unanticipated context of application and even presentational aspects emerging from search engines and social media.

The ECPAIS suites of ethics certification criteria are currently being extended to include ethical Privacy and tailored suites for high social impact domains including a bespoke suite for ethical assurance of COVID-19 pandemic related Contact Tracing Technologies [12]. This trend will continue to ensure ECPAIS embodies a broader and more comprehensive range of concerns in technology ethics.

6. Conclusions and the way forward

2021 should also be the age of AI ethics implementation, where operationalising AI ethics is not only seen as building trust to secure your customer base, an innovation enabler, providing legitimacy, and/or a competitive advantage, but an opportunity to build back better, recognising and addressing systemic inequalities and injustices, and creating a level playing field for people no matter who they are, their socio-economic circumstances, their background, or where they are in the world.

We ought not to consider AI and its application to the world in terms of an unlevel playing field but rather the world is the field in which everyone must play, how can we all work together best to create a playing field that everyone so that ALL can survive, be dignified and respected, thrive and flourish in using AI, with no one left behind [13].

In this endeavour, we ought to recognise that after two millennia of recorded civilization, consideration of ethics and social values in all that we do is a long overdue development. This therefore is a journey that thanks to the emergence of ADM/ALS we have just embarked on and should not be treated as a destination in line with many other facets and emergent properties of products, services and systems.

IntechOpen

Author details

Ali Hessami^{1*} and Patricia Shaw²

1 Vega Systems, United Kingdom

2 Beyond Reach – A Tech Ethics (Governance, Policy and Legal) Consultancy, United Kingdom

*Address all correspondence to: hessami@vegaglobalsystems.com

IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] UN Resolution. The Right to Privacy in a Digital Age, 26th September 2019, <https://digitallibrary.un.org/record/3837297?ln=en>
- [2] The EU has sought after creating an ecosystem of excellence and of trustworthy AI: https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence_en
- [3] Human Rights Council. The right to privacy in the digital age, Annual report of the United Nations High Commissioner for Human Rights and reports of the Office of the High Commissioner and the Secretary-General, A/HRC/39/29
- [4] https://www.vanguardinvestor.co.uk/investing-explained/ESG-funds?cmpgn=PS0820UKBABES0001EN&skwcid=AL!11156!3!471139682346!e!g!vanguard%20esg%20funds&gclid=CjwKCAjwo4mIBhBsEiwAKgzXOLNIh1QevqqZ3sSHICDi rehFI5ZahQn_rOsFYJGHTW3xqeEnNaaXhoCUTsQAvD_BwE&gclsrc=aw.ds
- [5] <https://www.theguardian.com/money/2020/jun/13/ethical-investments-are-outperforming-traditional-funds>
- [6] Bekhradnia, S. (2007). The Beliefs of Zoroastrianism, *New Statesman*, 8 January 2007
- [7] Hessami, A. G. Perspectives on Risk, Assessment and Management Paradigms, Intech Open, ISBN 978-1-83880-134-2, 2019
- [8] Examining the Black Box: Tools for assessing algorithmic systems, Ada Lovelace Institute, 29th April 2020, <https://www.adalovelaceinstitute.org/report/examining-the-black-box-tools-for-assessing-algorithmic-systems/>
- [9] Algorithmic accountability for the public sector, Ada Lovelace Institute, dated 24th March 2021, <https://www.adalovelaceinstitute.org/project/algorithmic-accountability-public-sector/>
- [10] Exploring the UK's digital divide – Office for National Statistics, dated 4th March 2019 <https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/articles/exploringtheuksdigitaldivide/2019-03-04#the-scale-of-digital-exclusion-in-the-uk>
- [11] <https://standards.ieee.org/industry-connections/ecpais.html>
- [12] IEEE Use Case—Criteria For Addressing Ethical Challenges In Transparency, Accountability, And Privacy Of Contact Tracing—Call For Global Consultation. <https://beyondstandards.ieee.org/new-creative-commons-paper-addresses-ethical-hurdles-to-contact-tracing-adoption/>
- [13] Hutter, B. M. (2005). The Attractions of Risk-based Regulation: accounting for the emergence of risk ideas in regulation, ESRC Centre for Analysis of Risk and Regulation, Discussion Paper 33, London School of Economics & Political Science March 2005