

Forum: Legal Committee (GA6)

Issue: Ensuring accountability and transparency in the use of algorithms, especially in decision-making processes

Student Officer: Maria Artopoulou

Position: Chair

INTRODUCTION

Algorithms define our 21st-century world, from selecting content for social media feeds to identifying diseases on X-rays. Modern algorithms like GPT are transforming the world by enabling computers to accomplish activities more efficiently than humans and even outperforming them in certain cases. However, it is wrong to assume that algorithms are recent creations since they have been used for thousands of years, even when not labelled as such. What has changed over the years are the types of algorithms and how they are used. A prime example of modern algorithm usage is decision-making, which uses most frequently machine-learning technologies that offer the possibility to prevent the drawbacks of human biases and make the process quicker. Hence, many fields have adopted automated decision systems, including employment companies, healthcare, justice, police, and banking. Nevertheless, the data stored in these systems can be already tainted from prejudice, by the learning process itself or its programmer, unavoidably leading to the creation of discriminatory patterns. Therefore, it can backfire and instead of limiting biased decisions, discriminating even further against certain groups. The President of the General Assembly Francis Dennis himself said at the gathering of the General Assembly that “No one should be at the mercy of an algorithm they don’t control, which was not designed to safeguard their interests, and which tracks their behaviour to collect personal data and keep them hooked”. At the same gathering, the Assembly adopted the first draft resolution ever regarding the Use of Artificial Intelligence Systems without a vote. This milestone resolution is the most recent act to regulate such systems that may uncontrollably breach multiple fundamental human rights and highlights the importance of setting regulations for safe AI use.

DEFINITION OF KEY-TERMS

Accountability

“Accountability is a principle according to which a person or institution is responsible for a set of duties and can be required to give an account of their fulfilment to an authority in a position

to issue rewards or punishment”¹. When referring to algorithms, it means having the ability to hold developers and users accountable for the results that these systems produce.

Algorithm

“An algorithm is a set of mathematical instructions that must be followed in a fixed order, and that, especially if given to a computer, will help to calculate an answer to a problem”². They are the foundation of automated decision-making processes, allowing systems to handle complicated duties.

Artificial intelligence

“the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings. The term is frequently applied to the project of developing systems endowed with the intellectual processes characteristic of humans, such as the ability to reason, discover meaning, generalize, learn from experience”³, and make decisions.

Bias

“The fact that a collection of data has more material that supports a specific perspective than you would expect to discover if the collection was generated by chance”⁴

Black box algorithm

“A black box algorithm is one where the user cannot see the inner workings of the algorithm. It is a rather controversial system, due to the secrecy they contain and the lack of transparency, although its creators defend it as a security and privacy system to avoid data leaks and unfair competition.”⁵

Decision making progress

The decision-making process is the methodical procedure for locating and assessing possibilities to make a decision. It entails a number of stages intended to collect and evaluate data, weigh options, and decide on the best course of action.

Heuristics

¹ Castiglione, Dario. “Accountability | Social Science | Britannica.” *Encyclopædia Britannica*, 2019, www.britannica.com/topic/accountability.

² Cambridge Dictionary. “ALGORITHM | Meaning in the Cambridge English Dictionary.” *Cambridge.org*, 2019, www.dictionary.cambridge.org/dictionary/english/algorithm.

³ Copeland, B.J. “Artificial Intelligence.” *Encyclopedia Britannica*, 11 Sept. 2022, www.britannica.com/technology/artificial-intelligence.

⁴ ---. “Bias.” *Cambridge.org*, 2019, www.dictionary.cambridge.org/dictionary/english/bias.

⁵ “What Is Black Box Algorithm - Definition and Examples.” *Armetrics*, 16 Nov. 2021, www.armetrics.com/en/digital-glossary/black-box-algorithm.

“Involving or serving as an aid to learning, discovery, or problem-solving by experimental and especially trial-and-error methods or relating to exploratory problem-solving techniques that utilize self-educating techniques to improve performance”⁶

Transparency

“A situation in which activities are done openly without disclosing information so that the decisions are impartial and fair”⁷.

BACKGROUND INFORMATION

History of algorithms

The name algorithm originates from Abdullah Muhammad bin Musa al-Khwarizmi, a mathematician from the 9th century who is known to be "The Father of Algebra" and an extraordinary mind. However, algorithms were developed and utilised long before they were given a label. Ancient algorithms were primarily concerned with mathematical ideas linked to geometry and arithmetic, with a particular emphasis on the mapping of land, engineering, trade, and astrology. A number of significant algorithms were created during this period, such as Euclidean Algorithm, Pythagorean Theorem and Archimedes' Method with the value of pi. Furthermore, there were several notable advances throughout the Islamic Golden Age (8th to 14th century), especially in the areas of astronomy, mathematics, medicine, and philosophy. Algebra was advanced by the development of techniques for solving linear and quadratic equations. Later in time the period of the Renaissance and Enlightenment saw intellectual advancements that lay the foundation for current algorithmic thinking, even though the name "algorithm" was not used at the time. It was due to the focus on systematic approaches, rigorous mathematics, and rationality that algorithms were developed as formalised processes for handling technical difficulties. In the last three centuries, advances in computer science, technology, and interdisciplinary study have led to the widespread use of algorithms in a variety of fields. The way we live, work, and communicate has changed significantly as a result of the development of algorithms, as has the way we make our decisions.

Types of Algorithms for decision making

Optimisation and Heuristics

It is formulating a problem with choice variables in order to maximise or minimise an objective function under constraints that mirror the intricacies of the actual world. For issues that can be mathematically abstracted, optimum solutions can be identified; however, the method has difficulty solving problems that are too big or complicated to be optimal. On the other hand, Heuristics provide an algorithmic sequence of steps to identify workable solutions, frequently

⁶ “Definition of HEURISTIC.” *Merriam-Webster.com*, 2017, www.merriam-webster.com/dictionary/heuristic.

⁷ ---. “TRANSPARENCY | Meaning in the Cambridge English Dictionary.” *Dictionary.cambridge.org*, 2021, www.dictionary.cambridge.org/dictionary/english/transparency.

scaling well and improving repeatedly; but they do not ensure optimality and do not provide a means of gauging a solution's proximity to optimality.

Machine Learning (ML)

Machine Learning (ML) is a broad term that refers to a set of algorithms that may be used to learn from previous data, which makes it a useful tool for decision-making. ML can be descriptive, predictive, or prescriptive, in contrast to optimisation and heuristics, which are only prescriptive. Because machine learning (ML) algorithms can extract patterns and insights from enormous datasets, they are especially useful for generating informed judgements and accurate predictions. Supervised learning algorithms, for instance, can help with inventory management, consumer behaviour research, and risk assessment by forecasting future patterns based on existing data. A prescriptive type of machine learning called reinforcement learning can improve decision-making by teaching it the best course of action through trial and error in changing circumstances. Since machine learning is flexible and scalable, it is an essential tool for today's decision-making, when gaining insights from data is essential to preserving competitive advantages and improving operational effectiveness. Prime examples of the use of such algorithms are autonomous vehicles. Programmed to reach a set destination by the safest and quickest route, they learn from previous data to improve its route taking into consideration a variety of data, including traffic lights, weather conditions, turns on the road, and any possible obstacles.

Applications of decision-making algorithms

Decision-making algorithms are found in multiple industries, aiming to enhance scalability and efficiency. First off, in healthcare, it is used to analyse patients' data and medical images, in order to identify abnormalities and diseases. The algorithms used have the ability to make personalised treatment plans, by figuring out which medications are necessary and which to avoid, which all in all improves the outcomes while being more affordable.

In Finance, such algorithms are used to score credit, evaluating an individual's creditworthiness after thoroughly analysing their financial data. This ensures faster and more accurate approvals of loans. In addition, the algorithms play a key role in the detection of fraud, as they can easily identify transactions with unusual patterns, which could indicate fraudulent activities. In this way, consumers and financial institutions are protected.

In the justice system, they are used to assist in predictive policing, such as sufficient criminal hotspot prediction, which helps law enforcement allocate their resources more effectively. They also conduct risk assessments to evaluate the possibility of individuals reoffending, a tool useful for sentencing and parole sentencing. There are more fields that use such algorithms, like e-commerce and the car industry, to increase efficiency and profitability, but as in those aforementioned their use can also complicate the work in different ways.

Advantages

Incorporating algorithms in decision-making has a number of notable advantages. First of all, algorithms are capable of processing enormous volumes of data rapidly and precisely, offering insights and suggestions that are not achievable for people to complete manually. This effectiveness improves results across a range of industries, including marketing, finance, and healthcare, by fostering more data-driven and informed decision-making. Moreover, by tracking

changes in the environment and reacting to them faster and more precisely than people, algorithms increase safety. Through the use of anomaly detection, equipment failure prediction, and real-time operation adjustments, industrial automation lowers the possibility of human mistake and improves the general safety and dependability of industrial processes. In addition, algorithms have the ability to reveal hidden connections and patterns in data, allowing predictive analytics to foresee future patterns and behaviours. This skill is very helpful in fields like customer relationship management, customised medication, and fraud detection. Furthermore, organisations can manage growing amounts of data and transactions thanks to the scalability of algorithms, which minimises the need for additional human resources and maximise operational effectiveness and cost-effectiveness. This is especially important for companies that handle large databases, such as Google and Meta, because it helps boost their profitability by avoiding extra expenses and initiating faster decisions. Algorithms also help with real-time decision-making, which is important in situations like emergency response and stock trading where prompt decisions can have a big influence on results. Organisations may achieve a combination of speed, precision, and strategic insight by adding algorithmic help to human decision-making, which will spur innovation and give them a competitive edge.

Challenges and Ethical Considerations

There are a number of challenges and ethical concerns with using algorithms in decision-making. First and foremost, bias is a significant issue as algorithms may unintentionally reinforce preexisting prejudices in the data they are trained on, producing unfair or discriminating results. Furthermore, there is a lack of accessibility since a lot of algorithms function as "black boxes," with intricate, opaque workings that are hard to understand. When judgments made by these systems are difficult to comprehend or challenge, this opacity can erode confidence and accountability. Since algorithms frequently demand large volumes of personal data, there is an increased potential for abuse or unauthorised access. This raises ethical considerations surrounding privacy. To be more precise, when algorithms are used in decision-making processes, large amounts of personal data must be gathered, typically including private information like financing, health records, and identification.

Significant ethical and legal issues are brought up by this extensive data collecting, especially in relation to consent. People might not be aware of how much data is being collected or what will be done with it, which might result in their privacy to be violated. Moreover, storing such large information presents significant security issues since they are easy targets for information theft and cyberattacks. Re-identification attacks have the potential to disclose persons from datasets that appear to have been anonymised, despite the fact that anonymisation procedures are not always effective in safeguarding identities. Furthermore, the invasion of privacy may be exacerbated if data initially gathered for one reason is repurposed without express authorisation for purposes like targeted advertising or sale to other parties. These concerns are heightened by the use of algorithms in surveillance since widespread monitoring might violate civil rights. Moreover, entrusting crucial choices to algorithms may reduce human agency and accountability, creating moral conundrums regarding the necessary level of human supervision. Regulatory monitoring, strict ethical guidelines, openness requirements, and technical innovation must all be carefully balanced in order to address these concerns and guarantee that the use of algorithms in decision-making processes is just, responsible, and respectful of people's rights.

Human rights infringe

A number of human rights, including the right to privacy, the right to non-discrimination, the right to due process, and the right to work and fair labour practices, may be violated by the use of algorithms in decision-making processes. The right to privacy guaranteed by Article 12 of the Universal Declaration of Human Rights (UDHR) and Article 17 of the International Covenant on Civil and Political Rights (ICCPR) is breached by inadequate safeguarding in the processing of large amounts of personal data can, which results in unauthorised data collection and surveillance.

The right to non-discrimination, safeguarded by Article 7 of the UDHR and Article 26 of the ICCPR, is jeopardised when biased algorithms reinforce preexisting societal prejudices, leading to unfair treatment in areas like employment, financing, and policing. Predictive police algorithms, for instance, have come under criticism for unfairly singling out minority populations as their targets.

Additionally, when algorithms make choices without accountability or transparency, the right to due process—which is protected by Article 10 of the UDHR and Article 14 of the ICCPR—may be compromised. People may not be able to comprehend or contest choices that negatively impact them, such as those pertaining to credit scores or court rulings, as a result of this lack of transparency.

Lastly, the right to work and fair labor practices emphasized in Article 23 of the UDHR can be violated by hiring algorithms that have the potential to unjustly reject competent applicants on the basis of inaccurate data, undermining their right to equal employment chances. These violations show how important it is to have strong legal frameworks and supervision procedures in place to guarantee that the use of algorithms respects and upholds fundamental human rights.

MAJOR COUNTRIES AND ORGANISATIONS INVOLVED

Canada

In 2019, the Canadian government announced the Directive on Automated Decision-Making, requiring all federal organisations that use automated decision systems to guarantee the accountability, fairness, and transparency of their choices. Algorithmic Impact Assessments (AIAs) are mandated by this regulation in order to assess and reduce the risks related to prejudice and discrimination in AI systems. It also mandates that these systems be auditable and able to justify actions that have an impact on specific people. Furthermore, the Canadian Office of the Privacy Commissioner has taken an active role in advising and supervising the moral use of AI, stressing the value of respecting human rights and safeguarding privacy.

China

The Cyberspace Administration of China (CAC) has released laws in recent years mandating that businesses reveal the fundamental ideas, intended uses, and working processes of their algorithms. These guidelines, which are a component of the more comprehensive "Internet Information Service Algorithmic Recommendation Management Provisions," require algorithmic

systems to be transparent and to adequately convey the effects they have on users. In order to combat prejudice, bias, and abuse, the laws also require frequent inspections and reviews. To increase user control and transparency, businesses must provide customers the ability to opt out of algorithmic recommendation services. China hopes to safeguard consumer rights and guarantee the ethical and fair operation of algorithmic systems by implementing these policies, which are consistent with the government's overarching objectives of digital governance and data security. China's dedication to addressing the moral ramifications of AI and algorithmic technologies is demonstrated by these measures.

United Kingdom (UK)

A key player in the endeavour to ensure transparency and accountability in the use of algorithms is the UK government's Centre for Data Ethics and Innovation (CDEI), which conducts studies of algorithmic decision-making and offers recommendations on moral AI practices. The "Review into Bias in Algorithmic Decision-Making" paper from the CDEI offers suggestions for enhancing accountability and transparency, including conducting regular audits and impact analyses. Additionally, the "AI Auditing Framework," which offers comprehensive standards for organizations to guarantee that AI systems are transparent, equitable, and responsible, was released by the Information Commissioner's Office (ICO). In line with GDPR, the UK has passed the Data Protection Act of 2018, which gives people the right to explanations for choices made by automated systems.

European Union (EU)

In order to guarantee transparency and accountability in the application of algorithms, especially in decision-making processes, the European Union has taken a number of noteworthy steps. The General Data Protection Regulation (GDPR), which requires openness in automated decision-making and profiling, is the cornerstone of these initiatives. In order to promote transparency, people are entitled to know the reasoning behind, importance of, and effects of automated decisions that impact them under GDPR. Furthermore, the Artificial Intelligence Act that the EU is proposing would govern AI by classifying systems according to risk categories and placing strict limitations on applications that pose a high danger. In order to maintain responsibility, this entails duties for risk assessment, documentation, and human monitoring. These rules are intended to shield people against judgements made by AI that are discriminating, or biased. The EU wants to build a reliable AI ecosystem that respects human rights, fosters justice, and increases public trust in digital technology. To this end, it is imposing strict transparency and accountability regulations.

Institute of Electrical and Electronics Engineers (IEEE)

Through its Ethically Aligned Design project and the creation of international standards, the Institute of Electrical and Electronics Engineers (IEEE) has made substantial efforts to guarantee accountability and transparency in the use of algorithms, particularly in decision-making processes. The work of the IEEE is concentrated on developing standards and recommendations that support moral issues in AI and autonomous systems. For instance, the IEEE P7000 series has standards for algorithmic bias concerns (P7003) and autonomous system transparency (P7001). By offering transparency into the decision-making process and facilitating consequence tracking, these standards seek to guarantee that AI systems are built with explicit

accountability procedures. The IEEE works to reduce prejudices and improve design ethics by supporting developers' and users' sense of responsibility and ethical design standards.

Organization for Economic Cooperation and Development (OECD)

The OECD adopted in 2019, Principles on AI provide guidelines to encourage the appropriate application of AI. Making sure algorithms are used responsibly and transparently—especially when it comes to decision-making processes—is one of the most important things to accomplish. These guidelines support accessibility and comprehensibility in AI systems so that people may understand and question AI-driven choices that impact them. The guidelines place a strong emphasis on the need for transparent documentation and evaluation procedures to track the effects of AI and guarantee that the systems function in an impartial and moral manner. By making companies and individuals accountable for the results of AI systems and making sure they abide by the law and set ethical standards, accountability is enforced. This strategy seeks to increase public confidence in AI technology while reducing the likelihood of prejudice and discrimination.

World Economic Forum (WEF)

Through its programmes and frameworks, the World Economic Forum (WEF) has taken the lead in promoting accountability and openness in the use of algorithms, particularly in decision-making processes. In addition to creating the Global AI Council, the WEF's Centre for the Fourth Industrial Revolution has released papers with titles like "AI Governance: A Holistic Approach to Implement Ethics into AI" and "Empowering AI Leadership." These publications offer thorough methods that prioritise transparency and accountability for incorporating ethical issues into AI research and use. In order to guarantee that judgements made by AI are just and transparent, the WEF encourages the implementation of AI ethical guidelines, which include thorough documentation of AI systems, frequent audits, and stakeholder participation. The WEF also helps multi-stakeholder partnerships to develop strong governance structures that comply with international standards.

TIMELINE OF EVENTS

18 th December 2013	The UN resolution on the Right to Privacy in the Digital Age was adopted to address growing concerns over privacy violations due to digital technologies
23 rd February 2018	The first FAT/ML Conference is held
25 th May 2018	The General Data Protection Regulation (GDPR) is enacted by the EU
22 nd May 2019	The OECD Principles are adopted
12 th January 2021	The Information Commissioner's Office in the UK releases guidelines for ensuring AI

	systems are transparent, fair, and accountable, enhancing GDPR's principles
27 th August 2021	China's Internet Information Service Algorithmic Recommendation Management Provisions is released introducing laws requiring companies to disclose fundamental principles, intended uses, and operational processes of algorithms
18 th November 2021	The United Kingdom's CDEI Review into Bias in Algorithmic Decision-Making is released by the UK's Centre for Data Ethics and Innovation offering recommendations for improving transparency and accountability, including regular audits and impact analyses
25 th November 2021	UNESCO's Recommendation on the Ethics of Artificial Intelligence is adopted setting a global framework for ethical AI development and application, emphasizing transparency, accountability, and equity to prevent bias in algorithmic decision-making
3 rd February 2022	The United States proposes the Algorithmic Accountability Act, aiming to prevent discrimination and ensure ethical use of data
1 st March 2023	China's Expanded Algorithm Regulations are strengthened, mandating regular inspections and user control over algorithmic recommendation services, promoting transparency and fairness.
21 st May 2024	The UN Resolution A/78/L.49 adopted by the General Assembly emphasises the need for "safe, secure, and trustworthy" AI systems

RELEVANT UN RESOLUTIONS, TREATIES AND EVENTS

RESOLUTION A/78/L.49

On March 21, 2024, the UN General Assembly passed Resolution A/78/L.49, which was an important milestone towards regulating artificial intelligence (AI). "This resolution emphasizes the importance of developing "safe, secure, and trustworthy" AI systems that respect human rights



and contribute to sustainable development.”⁸ It seeks to guarantee that, especially in light of AI, the rights people have offline are upheld online. Because it addresses the possible drawbacks and advantages of AI technologies, this resolution is very significant. In order to reduce the digital gap and give developing nations access to and benefits from AI developments, it advocates for international cooperation. Additionally, it encourages the global community to set up regulations to guarantee AI systems don't infringe upon human rights or present excessive threats. Regarding algorithmic decision-making, the resolution aims to shield people from prejudice and discrimination that artificial intelligence (AI) systems might reinforce. Through encouraging open and responsible AI development, it strives to protect individuals' rights from unjust automated choices that may have a substantial impact on their lives. All things considered, the UN General Assembly's move towards AI governance is proactive and inclusive. It emphasises how international cooperation is necessary to control AI's societal impacts and make sure it's applied fairly and ethically. The resolution seeks to safeguard basic human rights while utilising AI for the benefit of all people by establishing these guidelines.

UNESCO's Recommendation on the Ethics of Artificial Intelligence (2021)

Adopted in 2021, UNESCO's Recommendation on the Ethics of Artificial Intelligence offers a worldwide framework to direct the moral advancement and application of AI. In order to guarantee that AI technologies are developed and applied in a way that upholds human rights, democratic ideals, and the rule of law, this suggestion sets out guiding principles and values. It focuses on openness, responsibility, and equity in AI systems, aiming to avoid prejudice and discrimination that may result from algorithmic decision-making. UNESCO's recommendations is an important step forward for global attempts to regulate AI morally. It offers a standard framework that nations may take up and modify to fit their unique circumstances, encouraging a concerted international response to AI ethics. It helps reduce risks and guarantees AI research is in line with social values by highlighting human rights and ethical issues, resulting in a more just and equitable technological environment.

UN Resolution on the Right to Privacy in the Digital Age (2013)

Adopted in 2013, the UN Resolution on the Right to Privacy in the Digital Age tackles the increasing concerns about privacy infringements brought on by digital technology and widespread surveillance. With a focus on privacy protection and personal data security, the resolution advocates that people have the same rights online as they have offline. In the absence of appropriate legislation, digital communication as well as data processing technologies have the potential to easily breach people's right to privacy. The resolution is especially important when it comes to algorithmic decision-making, because algorithms may have an influence on people's lives without always being transparent or accountable. Examples of these applications include financial services and enforcement of laws. In the digital era, safeguarding personal information is making sure that algorithms don't infringe people's rights through discriminatory practices, profiling, or unauthorised data usage. It is an essential step in the right direction to create global

⁸ United Nations. “General Assembly Adopts Landmark Resolution on Artificial Intelligence | UN News.” *News.un.org*, 21 Mar. 2024, www.news.un.org/en/story/2024/03/1147831.

standards and legal frameworks that protect privacy and encourage a balance between the development of technology and fundamental human rights.

PREVIOUS ATTEMPTS TO SOLVE THE ISSUE

General Data Protection Regulation (GDPR) (2018)

Enacted in 2018, the General Data Protection Regulation (GDPR) is a comprehensive EU law designed to safeguard privacy and personal data. It is applicable to every organisation, worldwide, that handles the personal data of citizens of the European Union. GDPR imposes rigorous standards on the gathering, usage, and storing of personal data. It also demands transparency and explicit consent from individuals. It holds significance because it gives individuals greater control over their confidential information and imposes severe consequences for violating the regulations. In particular, algorithmic decision-making is covered by GDPR, which guarantees transparency and equity by granting people the right to request human intervention and to be made aware of automated decisions that directly impact them. Although there have been complaints about the GDPR's complexity and implementation costs, overall, it represents a substantial advancement in data protection, encouraging higher standards of privacy and influencing global data security norms. It is an essential step towards bringing legal structures into line with the demands of the digital age.

Algorithmic Accountability Act

The Algorithmic Accountability Act proposed by the US seeks to address the effects of commercial usage of automated decision-making systems. Companies are required by this legislation to evaluate the security, privacy, fairness, and bias of their algorithms, especially those that have a substantial impact on choices affecting a consumer's credit, employment, or insurance. In order to make sure that these algorithms do not violate privacy rights or encourage discrimination, the law mandates frequent impact evaluations and transparency reports. This law is significant because it aims to reduce the dangers connected to algorithmic decision-making and guarantee the open and impartial operation of such systems. By requiring accountability mechanisms that stop prejudices and unfair treatment ingrained in automated systems, it directly relates to the protection of people's rights by preventing discrimination and guaranteeing the ethical use of data. Overall, the Algorithmic Accountability Act is a significant step towards responsible control of AI. The imposition of responsibility and transparency on algorithmic processes serves to safeguard individual rights and enhance fairness, all the while fostering a sense of confidence in technology progress. This legislative attempt emphasises the necessity of thorough regulation in the digital economy to guarantee that everyone is treated fairly.

Fairness, Accountability, and Transparency in Machine Learning (FAT/ML) Conference

An annual academic conference on the ethical aspects of artificial intelligence and machine learning is called the Fairness, Accountability, and Transparency in Machine Learning (FAT/ML) Conference. The conference brings together researchers, practitioners, and policymakers to discuss and establish best practices for designing algorithms that are transparent, responsible, and fair. It was founded in response to rising concerns about the societal consequences of AI and machine learning. This conference is notable because it tackles crucial topics that can have a big influence on people's lives, such as prejudice, discrimination, and lack of transparency in algorithmic decision-making. FAT/ML seeks to promote the creation of algorithms

that are not only technically sound but also socially and ethically sound by encouraging multidisciplinary cooperation and discourse. By giving ethical issues first priority, the FAT/ML Conference actively contributes to influencing the direction of AI and machine learning in the future. It encourages the development of human rights-abiding technology and builds confidence in automated systems.

Ethical guidelines for trustworthy AI

The European Commission's ethical guidelines for trustworthy AI, published in 2019, outline principles to ensure that artificial intelligence systems are developed and deployed in a manner that is ethical and beneficial for society. These guidelines focus on seven key requirements: “human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental well-being, and accountability.”⁹ These guidelines emphasise the need for AI to respect fundamental rights and adhere to laws, ensuring that AI systems are not only effective but also ethical. Furthermore, the focus on the betterment of society and the environment guarantees that the advancement of AI is in line with more general social objectives, such as social justice and sustainability. Nonetheless the execution of these directives poses obstacles, given that it necessitates collaboration among numerous stakeholders and constant oversight to accommodate technological progressions. All things considered, these rules offer a strong ethical foundation that may direct the responsible creation and application of AI.

POSSIBLE SOLUTIONS

Creation of Standards and certification processes

In order to guarantee accountability and transparency in the application of algorithms in decision-making processes, standards and certification procedures for algorithms must be developed. To standardise algorithms, a collection of best practices and accepted norms must be established. These standards should address issues like equity, bias reduction, data management, and decision-making openness. Independent evaluations should be conducted as part of certification procedures to confirm that an algorithm conforms with these requirements. In this effort, international organisations like the International Organisation for Standardisation (ISO) can be extremely important. The ISO has the capacity to create international standards that offer a uniform foundation for algorithmic accountability and transparency that are relevant to many nations and sectors. These guidelines would guarantee that algorithms are created, implemented, and overseen in a way that respects human rights and ethical norms. Internationally, these standards would promote confidence and collaboration across borders, allowing nations to enact and execute these policies consistently and lowering the possibility of divergent legislative strategies impeding the advancement of AI worldwide. Governments at the national level can incorporate these international standards into their regulatory frameworks to provide businesses and developers clear, standardised guidelines to follow. This would provide a clear and consistent regulatory framework that would encourage innovation while also increasing public trust in algorithmic decision-making. Algorithms would be evaluated against these criteria by certification

⁹ European Commission. “Ethics Guidelines for Trustworthy AI | Shaping Europe’s Digital Future.” *European Commission*, 8 Mar. 2021, digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

bodies recognised by national or worldwide standardisation organisations, guaranteeing that only systems that comply are implemented. This strategy is critical because it establishes a strong framework for holding algorithm creators responsible, guarantees openness in their business practices, and encourages the moral use of AI technology everywhere.

Designing Ethical Algorithm Principles

Adopting ethical design principles is crucial to guaranteeing the accountability, fairness, and transparency of algorithms utilised in decision-making processes. Developers can use these principles as a guide for creating algorithms that respect human rights and adhere to ethical norms. A focus on justice entails aggressively eradicating prejudices that can produce unfair results. Making the algorithmic decision-making process accessible and intelligible for stakeholders is the goal of transparency, while accountability makes sure that users and developers can be held accountable for the algorithm's effects. Including diverse teams in the design process is essential to putting these ideals into practice. Diverse teams provide a variety of viewpoints and life experiences. This diversity is crucial for making sure the algorithms are inclusive and equitable for all users.

Establishment of Legislations

One of the most important steps in guaranteeing the ethical and responsible application of AI is the implementation of legislation requiring openness and responsibility in the usage of algorithms. By enforcing standards for impact assessments, audits, and the right to explanation, legislation may provide a framework that holds algorithm creators and users responsible for their activities. Audits entail routinely and methodically checking algorithms to make sure they adhere to moral principles and don't yield unfair or biased results. Impact analyses analyse how an algorithm could affect various demographic groups, assisting in the detection and mitigation of any negative consequences prior to implementation. People have the right to an explanation because it allows them to know how choices that impact them were made, which promotes transparency and gives them the chance to question or reject these judgements.

BIBLIOGRAPHY

“Adisa, Dorcas. “Everything You Need to Know about Social Media Algorithms.” *Sprout Social*, 30 Oct. 2023, sproutsocial.com/insights/social-media-algorithms/.

“AI Watch: Global Regulatory Tracker - United Nations | White & Case LLP.

Www.whitecase.com, 13 May 2024, www.whitecase.com/insight-our-thinking/ai-watch-globalregulatory-tracker-united-nation s. Accessed 18 July 2024.

“AI’s Ethical Issues: Getting around the Moral Maze.” *Www.linkedin.com*,

www.linkedin.com/pulse/ais-ethical-issues-getting-around-moral-maze-techcurators-ucfte.

Accessed 31 July 2024.

Algorithmic Accountability Act of 2023.

“Algorithms for Decision Making: Optimization, Heuristics and Machine Learning.”

Www.linkedin.com, www.linkedin.com/pulse/algorithms-decision-making-optimization-heuristics-machine-potluri . Accessed 31 July 2024.

Asie, and Visions. *Cyberspace Governance in China Evolution, Features and Future Trends*.

No. 129, 2022,

www.ifri.org/sites/default/files/atoms/files/lee_cyberspace_governance_china_2022.pdf

Belenguer, Lorenzo. “AI Bias: Exploring Discriminatory Algorithmic Decision-Making Models and the Application of Possible Machine-Centric Solutions Adapted from the Pharmaceutical Industry.” *AI and Ethics*, vol. 2, no. 2, 10 Feb. 2022, link.springer.com/article/10.1007/s43681-022-00138-8, <https://doi.org/10.1007/s43681-022-00138-8>.

Burns, Ed, and Kate Brush. “What Is a Decision-Making Process?” *SearchBusinessAnalytics*, Mar. 2020, www.techtarget.com/searchbusinessanalytics/definition/decision-making-process.

Cambridge Dictionary. “ALGORITHM | Meaning in the Cambridge English Dictionary.”

Cambridge.org, 2019, dictionary.cambridge.org/dictionary/english/algorithm.

---. “Bias.” *Cambridge.org*, 2019, dictionary.cambridge.org/dictionary/english/bias.

---. “TRANSPARENCY | Meaning in the Cambridge English Dictionary.”

Dictionary.cambridge.org, 2021, dictionary.cambridge.org/dictionary/english/transparency.

Castiglione, Dario. “Accountability | Social Science | Britannica.” *Encyclopædia Britannica*, 2019, www.britannica.com/topic/accountability.

Copeland, B.J. “Artificial Intelligence.” *Encyclopedia Britannica*, 11 Sept. 2022,

www.britannica.com/technology/artificial-intelligence.

“Data & AI Insights: Empowering Decisions with Artificial Intelligence and Analytics.”

Www.linkedin.com, www.linkedin.com/pulse/data-ai-insights-empowering-decisions-artificial-intelligence-jha-1f.

“Definition of HEURISTIC.” *Merriam-Webster.com*, 2017,

www.merriamwebster.com/dictionary/heuristic.

Dieckmann, Jonas. "Ethics in AI: Potential Root Causes for Biased Algorithms." *Medium*, 21 Mar. 2023, towardsdatascience.com/ethics-in-ai-potential-root-causes-for-biased-algorithms-890091915aa3.

European Commission. "Ethics Guidelines for Trustworthy AI | Shaping Europe's Digital Future." *European Commission*, 8 Mar. 2021, digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

Government of Canada, Statistics Canada. "Responsible Use of Automated Decision Systems in the Federal Government." *Www.statcan.gc.ca*, 7 June 2021, www.statcan.gc.ca/en/data-science/network/automated-systems.

Hesmondhalgh, David, et al. "The Impact of Algorithmically Driven Recommendation Systems on Music Consumption and Production - a Literature Review." *GOV.UK*, 9 Feb. 2023, www.gov.uk/government/publications/research-into-the-impact-of-streaming-services-algorithms-on-music-consumption/the-impact-of-algorithmically-driven-recommendation-systems-on-music-consumption-and-production-a-literature-review.

"Home :: FAT ML." *Www.fatml.org*, www.fatml.org/.

"How Can Algorithms Aid in Human Decision-Making Processes?" *Www.linkedin.com*, www.linkedin.com/pulse/how-can-algorithms-aid-human-decision-making . Accessed 31 July 2024.

Jayatilake, Senerath Mudalige Don Alexis Chinthaka, and Gamage Upeksha Ganegoda.

"Involvement of Machine Learning Tools in Healthcare Decision Making." *Journal of Healthcare Engineering*, vol. 2021, 27 Jan. 2021, pp. 1–20, <https://doi.org/10.1155/2021/6679512>.

Platform, European Liberties. "Are Decision-Making Algorithms Always Right, Fair and Reliable or NOT?" *Liberties.eu*, 25 Mar. 2022, www.liberties.eu/en/stories/decision-making-algorithm/44109.

Ramnani, Shweta. *Exploring Ethical Considerations of Artificial Intelligence in Educational Settings: An Examination of Bias, Privacy, and Accountability*. Vol. 9, 2024, pp. 2456–4184, www.ijnrd.org/papers/IJNRD2402124.pdf.

Rodgers, Waymond, et al. “An Artificial Intelligence Algorithmic Approach to Ethical Decision-Making in Human Resource Management Processes.” *Human Resource Management Review*, vol. 33, no. 1, June 2022, p. 100925, www.sciencedirect.com/science/article/pii/S1053482222000432, <https://doi.org/10.1016/j.hrmr.2022.100925>.

Secretariat, Treasury Board of Canada. “Directive on Automated Decision-Making.” *Www.tbs-sct.canada.ca*, 5 Feb. 2019, www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592.

Sess.: 2020-2021), UN General Assembly (75th. “The Right to Privacy in the Digital Age :: Resolution /: Adopted by the General Assembly.” *Digitallibrary.un.org*, 28 Dec. 2020, digitallibrary.un.org/record/3896430.

Taxonomy of Human Rights Risks Connected to Generative AI Supplement to B-Tech’s Foundational Paper on the Responsible Development and Deployment of Generative AI.

Teja, Mohan. “History of Algorithms.” *Medium*, 2 Mar. 2024, medium.com/@mohantejadharmavarapu1997/history-of-algorithms-3562e169c33e.

“The Right to Due Process.” *Icelandic Human Rights Centre*, www.humanrights.is/en/human-rights-education-project/human-rights-concepts-ideas-and-fora/substantive-human-rights/the-right-to-due-process.

The Upwork Team. “How AI Is Used in Decision-Making Processes | Upwork.” *Www.upwork.com*, 16 Aug. 2023, www.upwork.com/resources/ai-in-decision-making.

Tsamados, Andreas, et al. “The Ethics of Algorithms: Key Problems and Solutions.” *AI & Society*, vol. 37, no. 1, 20 Feb. 2021, pp. 215–230, link.springer.com/article/10.1007/s00146-021-01154-8, <https://doi.org/10.1007/s00146-021-01154-8>.

United Nations. “General Assembly Adopts Landmark Resolution on Artificial Intelligence | UN News.” *News.un.org*, 21 Mar. 2024, news.un.org/en/story/2024/03/1147831.

---. “United Nations Charter.” *United Nations*, 1945, www.un.org/en/about-us/un-charter/full-text.

“What Is Black Box Algorithm - Definition and Examples.” *Arimetrics*, 16 Nov. 2021, www.arimetrics.com/en/digital-glossary/black-box-algorithm.

Wolford, Ben. “What Is GDPR, the Eu’s New Data Protection Law?” *GDPR.eu*, 2020, gdpr.eu/what-is-gdpr/.

Zhang, Renda. “Optimization Theory Series: 9 — Optimization Algorithms.” *Medium*, Medium, 11 Jan. 2024, rendazhang.medium.com/optimization-theory-series-9-optimization-algorithms-f381917c9a4a. Accessed 31 July 2024.