Ética na Inteligência Artificial: como apoiar os desenvolvedores?

Geber L. Ramalho

Centro de Informática – Universidade Federal de Pernambuco (UFPE)
Av. Jornalista Aníbal Fernandes, s/n – Cidade Universitária – Recife - PE – Brasil

glr@cin.ufpe.br

***Abstract.** This article examines the primary ethical challenges associated with Artificial Intelligence (AI), particularly in the development of Machine Learning systems. Although ethical principles have been increasingly consolidated, a significant gap remains between these principles and their practical implementation by developers. The study reviews the state of the art and emphasizes the "ethics by design" approach, underscoring the need to rethink the training of professionals in the field, as well as the development of tools, processes, and best practices aligned with their day-to-day activities.*

***Resumo.** Este artigo discute os principais desafios éticos da Inteligência Artificial (IA), especialmente no desenvolvimento de sistemas de Aprendizado de Máquina. Apesar da consolidação de princípios éticos, persiste uma lacuna entre esses princípios e sua aplicação prática por desenvolvedores. O texto analisa o estado da arte e destaca a abordagem "ethics by design", que evidencia a necessidade de repensar a formação desses profissionais, bem como o desenvolvimento de ferramentas, processos e boas práticas orientadas para suas atividades cotidianas.*

1. Introdução

A Inteligência Artificial (IA) tem emergido como uma força transformadora na sociedade moderna, trazendo revoluções em áreas tão diversas quanto saúde, finanças, educação e indústria. No entanto, juntamente com seus benefícios notáveis, a IA também traz uma série de questões éticas que precisam ser enfrentadas [Crawford 2021; Huang et al. 2022]. Enfrentá-las é fundamental para assegurar que a IA seja desenvolvida e utilizada de maneira que respeite os valores humanos e sociais, evitando consequências indesejadas. Há um consenso hoje sobre os princípios éticos da IA, mas *"the gap between principles and practice is large, and widened by complexity, variability, subjectivity, and lack of standardization"* [Morley et al. 2020].

Apesar dos avanços em termos de leis e normas para traduzir os princípios em instrumentos práticos, o **desafio** que permanece aberto é **como formar e apoiar os profissionais de computação a criarem sistemas de IA que adiram a princípios éticos em todo seu ciclo de vida, do design ao uso, passando pelo desenvolvimento e implantação?** Este é o desafio por excelência atualmente dentro da área de ética na IA, que cresceu muito nos últimos 4 anos, junto com a disseminação dos sistemas de IA. É um desafio crucial para garantir que as inovações tecnológicas promovam o bem-estar social e não amplifiquem problemas socioeconômicos existentes.

2. Questões éticas da IA

No desenvolvimento e uso de sistemas de IA, em particular os de Aprendizagem de Máquina (AM), várias questões éticas emergem, e elas podem ser organizadas em três categorias principais.

Há questões relacionadas ao próprio funcionamento dos modelos de AM. A eficácia de um sistema de AM depende de sua confiabilidade técnica, da sua precisão em previsões e decisões. Falhas podem causar danos graves, especialmente em áreas críticas como veículos autônomos [Thadani et al. 2023]. A justiça é outra preocupação, pois os algoritmos de AM podem amplificar preconceitos presentes em dados históricos, perpetuando discriminação com base em raça, gênero ou status socioeconômico, como visto no sistema COMPAS [Kirkpatrick 2017]. A autonomia humana também não pode ser posta em risco com a crescente automação de decisões [Tiribelli 2023], e garantir o alinhamento dos objetivos das máquinas com os dos seres humanos problema complexo [Yudkowsky 2016]. Além disso, muitos modelos de AM, especialmente os de aprendizagem profunda, são consideradas caixas-pretas, o que dificulta a explicabilidade e transparência [Dwivedi 2023]. Questões de responsabilização também existem, como no caso dos veículos autônomos, onde a culpa por acidentes pode ser difícil de atribuir entre desenvolvedores e usuários [Mueller et al. 2021].

Os dados, insumos fundamentais para os sistemas de AM, levantam preocupações éticas sobre privacidade e segurança. A coleta, armazenamento e uso de dados pessoais devem proteger os direitos dos indivíduos [Aldboush et al. 2023]. A segurança dos dados contra vazamentos é crucial para manter a confiança nos sistemas de IA [Mothukuri et al. 2021]. A transparência sobre o uso dos dados também é importante para garantir maior responsabilização [McCoy et al. 2023]. Questões de direitos autorais estão em pauta quando dados protegidos são usados sem autorização [Chesterman 2024].

O impacto social dos sistemas de AM inclui preocupações com o desemprego devido à automação, a proteção dos direitos humanos [Ximenes, Salcedo & Ramalho

2020], e o impacto na democracia por meio da disseminação de desinformação e publicidade política direcionada [Łabuz & Nehring 2024]. Além disso, a pegada de carbono dos modelos de AM, especialmente os de aprendizado profundo, é significativa, exigindo novas estratégias para reduzir emissões [CNBC 2024]. Por fim, há potenciais riscos de longo prazo envolvendo sistemas superinteligentes em relação aos interesses humanos [Bengio, 2023].

3. Fundamentos éticos para IA

Não há uma definição precisa de IA ética, entretanto, para este documento vamos adotar que **um sistema de IA ético é aquele que segue princípios éticos em todo seu ciclo de vida: design, desenvolvimento, implantação e uso**. Cobrir todo o ciclo de vida é crucial, já que questões éticas podem surgir em qualquer etapa do desenvolvimento e impactar componentes além dos algoritmos e modelos de AM [Morley et al. 2020; Lu et al. 2024]. Por exemplo, um sistema de AM que usa dados protegidos sem autorização para treinar o modelo não deve ser considerado ético.

São cinco os princípios éticos para IA, amplamente aceitos na literatura e entre especialistas [Floridi et al. 2018; Jobin et al. 2019]: (1) Beneficência, que visa promover o bem-estar, preservar a dignidade e sustentar o planeta; (2) Não-maleficência, que determina que a ação (mesmo benéfica) deve causar o menor dano possível, abordando questões como privacidade, confiabilidade técnica e segurança; (3) Justiça, que promove a prosperidade e preserva a solidariedade, estabelecendo a equidade como uma condição fundamental; (4) Autonomia, que preserva o poder de decidir (ou não decidir) e equilibra o poder de decisão que mantemos para nós mesmos e o que delegamos aos agentes artificiais; e (5) transparência, que envolve explicabilidade, inteligibilidade e responsabilização, garantindo a compreensão e a responsabilidade pelos processos de decisão dos sistema de IA.

Embora haja um consenso geral sobre a importância desses princípios, a tradução prática deles em diretrizes operacionais para o desenvolvimento de IA ainda é um desafio significativo, sobretudo para os desenvolvedores. De fato, “*as consensus across the various published AI ethics principles is approached, a gap remains between high-level principles and practical techniques that can be readily adopted to design and develop responsible AI systems*” [Sanderson et al. 2023]

4. Estado da arte

Diante dos desafios éticos decorrentes da rápida adoção e desenvolvimento de sistemas de IA, diversas comunidades de pesquisa surgiram para estudar o tema. Essas comunidades estão organizadas sob as denominações de "IA responsável", "IA confiável", "IA ética" e "IA segura". Apesar de pequenas diferenças em suas abordagens, elas compartilham o objetivo comum de garantir que as tecnologias de IA beneficiem a sociedade sem causar danos.

As iniciativas para traduzir princípios gerais em mecanismos práticos podem ser classificadas em quatro vertentes, seguindo o paradigma das forças de regulação de Lessig [Lessig 2006].

Primeiro, nas leis, houve inúmeros avanços nos últimos anos. O Regulamento Geral de Proteção de Dados (GDPR) da União Europeia é um marco na regulamentação de dados pessoais, estabelecendo novos padrões que influenciaram outros países, como os do BRICS e da América Latina [Belli & Doneda 2023]. Recentemente, a Europa propôs o AI Act, um marco regulatório para garantir o uso seguro e ético da IA, classificando sistemas de acordo com seus níveis de risco [European Union 2024]. O Brasil, depois de lançar a LGPD, está trabalhando em um Projeto de Lei sobre IA no senado [Senado 2023].

Nas normas, segunda força regulatória, não-obrigatória, houve também avanços. A ISO 27000, focada em segurança da informação e privacidade, é uma das primeiras normas relevantes, mas há hoje 16 certificações ISO/IEC relacionadas a sistemas de IA éticos, incluindo ISO/IEC 24028 (Confiabilidade em IA), 24027 (Viés em IA), 38507 (Governança da IA por organizações) e 23894 (Gestão de riscos) [Janačković et al. 2024]. Há trabalhos acadêmicos relevantes, com inclusive participação de brasileiros, sobre governança da IA de forma geral [Gasser & Almeida 2017].

O mercado, terceira força, via boicotes, auto-regulação e regulação imposta a fornecedores, pode impactar na regulação como as lojas de aplicativos da Apple e Google Play que estabelecem regras para classificar jogos voltados para crianças e adolescentes [Google 2024]. Diversos manuais de ética internos às empresas também surgiram [Sony Global 2024; Microsoft 2024].

O problema é que as regulações citadas são ou abstratas, ou complexas ou focadas em gestores, de forma que os desenvolvedores (arquitetos, cientistas de dados, engenheiros de AM, engenheiros de dados, engenheiros de software e operadores) não sabem como proceder no seu dia a dia para construir sistemas de IA éticos [Sanderson 2023, Lucaj et al. 2023]. De fato, *“dozens of proposals for addressing ethical aspects of artificial intelligence (AI) have been published. However, many of them are too abstract for being easily translated into concrete designs for AI systems”* [Prem 2023].

Para tanto, precisamos entrar no que Lessig chama de regulação por arquitetura, traduzida hoje por *“ethics by design”* e que é particularmente poderosa, pois busca desenvolver sistemas que sejam éticos desde a concepção, em vez de depender de leis, normas ou forças de mercado para mitigar problemas posteriores, sobretudo porque leis e normas nunca chegam na velocidade do avanço da tecnologia. É nesse quesito que se encontram o desafio para a computação que propomos e que se traduz na questão: como formar e apoiar os profissionais de computação a criarem sistemas de IA que adiram a princípios éticos em todo seu

ciclo de vida, do design ao uso, passando pelo desenvolvimento e implantação? Dentro dessa força de regulação, podemos ver as iniciativas atuais em **quatro eixos: educação; ferramentas e técnicas; processos e boas práticas**. Infelizmente, ainda estamos em estágios preliminares, com pouca validação, e praticamente nenhuma padronização.

Sobre a educação, os desenvolvedores geralmente não recebem treinamento ético extensivo. Educar desenvolvedores de sistemas de AM em princípios éticos, ferramentas e melhores práticas é uma etapa essencial para que eles priorizem justiça, transparência e responsabilidade, e que antecipem e mitiguem possíveis problemas ao desenvolver sistemas. Iniciativas de educação em IA ética estão surgindo [Tubella 2024; Alam 2023], identificando objetivos de aprendizado concretos e casos de uso implementáveis, mas ainda são muito poucas e longe de uma padronização ou adoção ampla. No CIn-UFPE, temos oferecido uma disciplina de ética em IA para a pós-graduação desde 2019, e algumas outras correlatas na graduação, mas elas ainda não formam um corpo coerente e coeso que abordem todas as competências necessárias sobretudo para os alunos da graduação.

Quanto às ferramentas e técnicas, houve muitos avanços nos últimos anos particularmente em segurança de dados, não discriminação e explicabilidade. Novas ferramentas estão surgindo a cada dia para ajudar os desenvolvedores a criar sistemas de AM éticos. Recentemente, foi apresentada uma compilação dessas ferramentas por meio de uma revisão sistemática [Prem 2023]. Exemplos incluem o AI Fairness 360 da IBM [IBM 2024], que oferece métricas e algoritmos para detectar e mitigar vies em modelos. Ferramentas de explicabilidade, como o SHAP [SHAP 2024], que ajudam a entender e explicar as decisões de modelos complexos, melhorando a transparência. Além disso, técnicas de preservação de privacidade, como a privacidade diferencial [Hassan 2019] e o aprendizado federado [Li et al. 2020], estão sendo integradas aos sistemas de IA para proteger os dados dos usuários. Claro, essas ferramentas tratam apenas de alguns pontos no ciclo de vida de sistemas de IA.

O terceiro eixo é o dos processos. Existem processos para criar sistemas de AM, desde o clássico CRISP-DM. No entanto, não há um sequer processo maduro para ajudar os desenvolvedores a criar sistemas de AM ou IA éticos. Alguns frameworks e metodologias emergentes estão sendo propostos para atender a essa necessidade na indústria, como o da Google [Google 2024b], mas ainda estão em estágios iniciais de adoção. Esforços acadêmicos propuseram processos para o desenvolvimento ético de IA, como o de Hundt [Hundt 2024] e o de Prem [Prem 2023], mas nenhum deles foi validado ainda. Outra iniciativa, o "Trustworthy Development Process" [Hohma & Lütge 2023], é uma coleção de diretrizes para desenvolvedores, mas também tem limitações de validação.

Enfim, sobre boas práticas, há diversas proposições nesse sentido, já que é mais fácil começar por elas do que por processos abrangentes. Mas há vários

problemas: (a) poucas das recomendações podem ser consideradas boas práticas no sentido de “medidas acionáveis”. A maioria é de “recomendações concretas”, com a vantagem de serem menos abstratas do que os princípios, porém menos claras do que boas práticas [Madaio et al. 2020]; (b) a maioria é focada em governança da IA voltada para gestores das empresas e no máximo gerentes de projetos, mas não para desenvolvedores [Chancellor 2023; Baldassarre et al. 2024]; (c) a maioria não tem uma validação empírica ou qualitativa por especialistas [Zhang et al. 2021; Lu et al. 2024]. Uma das poucas proposições de boas práticas validada é a que desenvolvemos no CIn-UFPE anos atrás [Ximenez & Ramalho 2021] e que está para ser publicada com mais profundidade e amplitude após um período de implantação de mais de um ano em uma empresa de IA com mais de 35 milhões de usuários ativos mensais.

5. Possíveis métricas de avanço no desafio

O desafio proposto reside em **transformar os princípios éticos da IA em práticas concretas e aplicáveis para desenvolvedores**, desafio que pode ser abordado a partir de quatro eixos fundamentais: educação dos desenvolvedores, processos de desenvolvimento, ferramentas e técnicas, e recomendações concretas e boas práticas. Assim, podemos tentar definir métricas para o avanço nesse desafio. Claro, as métricas já são um desafio em si dada a atualidade e a velocidade da adoção de IA, assim como sua versatilidade de aplicação, mas algumas métricas possíveis são:

- Quantidade de empresas que adotam ferramentas, processos e boas práticas éticas para sistemas de IA, em todo seu ciclo de vida;
- Número de frameworks, processos, e boas práticas propostas e validadas solidamente;
- Processos de auditoria adotados e validados;
- Número de cursos e treinamentos disponíveis em ética para IA, que deve ser amplo.

6. Conclusões e conexão com os outros desafios

A ética na IA é um campo emergente e vital para o futuro da tecnologia. Apesar dos avanços realizados, ainda há uma lacuna significativa entre os princípios éticos e sua aplicação prática, em especial para desenvolvedores de sistemas IA que não sabem como aderir no dia a dia aos princípios éticos para IA, se é que conhecem esses princípios. A criação de diretrizes concretas e boas práticas, acompanhadas por ferramentas, técnicas, processos e métricas adequados e eficazes, é essencial para garantir que a IA desenvolvida beneficie a sociedade de maneira ampla e equitativa. Avançar nesse desafio requer uma abordagem abrangente, que inclua também educação dos desenvolvedores.

Além desses, há desafios mais profundos na agenda da pesquisa em ética na IA. O primeiro, de caráter epistemológico, envolve a necessidade de se estabelecer um diálogo permanente e estruturado entre as comunidades da computação e das

humanidades. Não é possível falar de antecipar ou mitigar impactos negativos na sociedade sem envolver os pesquisadores que estudam a sociedade. Nós da computação habitualmente conversamos somente com nós mesmos.

O segundo, de caráter ontológico, envolve repensar o próprio conceito de algoritmo. Em tese recente na UFPE, é defendida a ideia de que estamos em uma transição de “closed algorithms” para “open algorithms”, dos quais os sistemas de aprendizado de máquina fazem parte [Falcão 2025]. Os “algoritmos abertos” são menos “controláveis” por diversas razões, entre elas porque não são apenas código, mas um amálgama de código, dados, pessoas e interfaces. Não controlamos as pessoas e por consequência os dados que elas provêm. Reconhecer esse novo “animal” pode nos ajudar a melhorar e criar ferramentas, processos, métricas e práticas que levem em conta princípios éticos.

A proposta de desafio está fortemente conectada aos desafios tecnológicos e socioeconômicos atuais da SBC. Em termos de desafios tecnológicos, o tema “IA ética” dialoga com a “IA” em si, “ciência de dados” (porque dados são o principal insumo dos modelos de IA), “ciber-segurança” (também por conta do uso dados) e “computação sustentável” (porque o treinamento dos modelos consome muita energia). Quanto aos desafios socioeconômicos, a IA ética se relaciona com praticamente todos também: diretamente com os “aspectos éticos da computação”, mas também com o “combate à desinformação” (pois IA ajudar a gerar ou supervisionar esse tipo de conteúdo), o “fortalecimento da inovação e do empreendedorismo” (visto que a IA será motor de muitas inovações), e “impactos socioeconômicos” que são a principal motivação da proposição desse desafio.

7. Referências

- Alam, Ashraf. Developing a Curriculum for Ethical and Responsible AI: A University Course on Safety, Fairness, Privacy, and Ethics to Prepare Next Generation of AI Professionals. *Intelligent Communication Technologies and Virtual Mobile Networks*. Singapore: Springer Nature Singapore, 2023. 879-894
- Aldboush et al. Building Trust in Fintech: An Analysis of Ethical and Privacy Considerations in the Intersection of Big Data, AI, and Customer Trust. *International Journal of Financial Studies* 11.3 (2023): 90.
- Baldassarre et al. Fostering Human Rights in Responsible AI: A Systematic Review for Best Practices in Industry. *IEEE Transactions on Artificial Intelligence* (2024).
- Belli, L., and Doneda, D. Data protection in the BRICS countries: legal interoperability through innovative practices and convergence. *International Data Privacy Law* 13.1 (2023): 1-24.
- Bengio, Y. (2023). AI and catastrophic risk. *Journal of democracy*, 34(4), 111-121.
- Chancellor, S. (2023). Chancellor, Stevie. Toward practices for human-centered machine learning. *Communications of the ACM* 66.3 (2023): 78-85.
- Chesterman, Simon. Good models borrow, great models steal: intellectual property rights and generative AI. *Policy and Society* (2024): puae006.

- Crawford K. *Atlas of AI: power, politics, and the planetary costs of artificial intelligence*, 1st edn. (2021). Yale University Press
- CNBC - Consumer News and Business Channel. Google's carbon emissions surge nearly 50% due to AI energy demand (2024). Available in <https://www.cnbc.com/2024/07/02/googles-carbon-emissions-surge-nearly-50percent-due-to-ai-energy-demand.html>
- Dwivedi et al. Explainable AI (XAI): Core ideas, techniques, and solutions. *ACM Computing Surveys* 55.9 (2023): 1-33.
- European Union - The EU Artificial Intelligence Act. Available in <https://artificialintelligenceact.eu> (2024)
- Falcão, J. (2025). *Open Algorithms: an interdisciplinary inquiry of artificial intelligence systems*. Tese de doutorado. Centro de Informática – UFPE.
- Floridi L, Cowls J, Beltrametti M, et al (2018) AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds Mach* 28:689–707.
- Gasser, U. and Virgilio Almeida. A layered model for AI governance. *IEEE Internet Computing* 21.6 (2017): 58-62.
- Google. Google Play Families Policies. Available in <https://support.google.com/googleplay/android-developer/answer/9893335?hl=en&sjid=13527497315176022638-SA#1&2&3&4&5&6&7&8&9&zippy=%2Cexamples-of-common-violations>. (2024)
- Google. Responsible AI. Available in <https://cloud.google.com/responsible-ai>. (2024b)
- Hassan, M., Mubashir Rehmani, & Jinjun Chen. Differential privacy techniques for cyber physical systems: A survey. *IEEE Communications Surveys & Tutorials* 22.1 (2019): 746-789.
- Hohma, E. & Christoph Lütge. From trustworthy principles to a trustworthy development process: The need and elements of trusted development of AI systems. *AI* 4.4 (2023): 904-925.
- Huang et al. An overview of artificial intelligence ethics. *IEEE Transactions on Artificial Intelligence* 4.4 (2022): 799-819.
- Hundt, A., Schuller, J. & Severin Kacianka. Towards Equitable Agile Research and Development of AI and Robotics. arXiv preprint arXiv:2402.08242 (2024).
- IBM. AI Fairness 360. Available in <https://aif360.res.ibm.com> (2024)
- Janačković, G., Vasović, D. and Bojan Vasović. Artificial Intelligence standardisation efforts. *Engineering management and competitiveness (EMC)* 2024 (2024): 250.
- Jobin et al. The global landscape of AI ethics guidelines. *Nat Mach Intell* 1:389–399. (2019)
- Kirkpatrick, K. It's not the algorithm, it's the data. *Communications of ACM* 60:21–23. (2017)
- Łabuz, M. & Nehring, C. On the way to deep fake democracy? Deep fakes in election campaigns in 2023. *European Political Science* (2024): 1-20.
- Lessig L *Code: And Other Laws of Cyberspace*, Version 2.0, 2nd edn. Basic Books (2006)
- Li, Tian, et al. Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine* 37.3 (2020): 50-60.
- Lu et al. Responsible AI pattern catalogue: A collection of best practices for AI governance and engineering. *ACM Computing Surveys* 56.7 (2024): 1-35.

- Lucaj, L., Van Der Smagt, P. and Djalel Benbouzid. "AI regulation is (not) all you need." Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. (2023).
- Madaio et al. Co-designing checklists to understand organizational challenges and opportunities around fairness in AI. Proceedings of the 2020 CHI conference on human factors in computing systems. (2020)
- McCoy et al. Ethical responsibilities for companies that process personal data. *The American Journal of Bioethics* 23.11 (2023): 11-23.
- Microsoft. Empowering responsible AI practices. Available in <https://www.microsoft.com/en-us/ai/responsible-ai>. (2024)
- Morley, J., Floridi, L., et al. From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. *Sci. Eng. Ethics* 26, 2141–2168 (2020).
- Mothukuri et al. A survey on security and privacy of federated learning, *Future Generation Computer Systems*, vol. 115, pp. 619–640, (2021).
- Mueller et al. Addressing driver disengagement and proper system use: human factors recommendations for level 2 driving automation design. *Journal of cognitive engineering and decision making* 15.1 (2021): 3-27.
- Prem, E. From ethical AI frameworks to tools: a review of approaches. *AI and Ethics* 3.3 (2023): 699-716.
- Sanderson et al. AI Ethics Principles in Practice: Perspectives of Designers and Developers, *IEEE Transactions on Technology and Society*, pp. 1–1, (2023).
- Senado Federal Brasileiro. Projeto de Lei nº 2338. Disponível em <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233> (2023)
- SHAP. Available in <https://shap.readthedocs.io/en/latest>. (2024)
- Sony Global. Sony Group's Initiatives for Responsible AI. Available in https://www.sony.com/en/SonyInfo/sony_ai/responsible_ai.html. (2024)
- Thadani, Trisha, et al. Tesla drivers run Autopilot where it's not intended-with deadly consequences. *The Washington Post* (2023): NA-NA.
- Tiribelli, S. The AI ethics principle of autonomy in health recommender systems. *Argumenta* 16 (2023): 1-18.
- Tubella, A. Marçal Mora-Cantalops, A. and Juan Carlos Nieves. How to teach responsible AI in Higher Education: challenges and opportunities. *Ethics and Information Technology* 26.1 (2024): 3.
- Ximenes, B. Salcedo, D. & Ramalho, G. Towards broadening the perspective on lethal autonomous weapon systems ethics and regulations. *Rio Seminar on Autonomous Weapons Systems*. (Rio de Janeiro, Naval War College) - Brasília: FUNAG, 2020
- Ximenes B & Ramalho G. Concrete Ethical Guidelines and Best Practices in Machine Learning Development. In: *Proceedings of IEEE International Symposium on Technology and Society*. (2021)
- Yudkowsky, Eliezer. "The AI alignment problem: why it is hard, and where to start." *Symbolic Systems Distinguished Speaker* 4.1 (2016).
- Zhang, T., Qin, Y., & Li, Q. (2021). Trusted artificial intelligence: technique requirements and best practices. *International Conference on Cyberworlds (CW)*. IEEE (2021).