

A whitepaper on

Creating AI Responsibly

Unlocking AI's Potential



October 2024
10-13-2024

Preface

5G India Forum (5GIF) is an industry collaborative forum under the aegis of Cellular Operators Association of India (an MRP of 3GPP) and was established to enable synergizing the national efforts and play a significant role in shaping the strategic, commercial and regulatory development of the 3GPP ecosystem in India. 5GIF plays a significant role in shaping the future of 5G and next generation 3GPP technologies in India by bringing together various stakeholders and working towards common goals in this rapidly evolving technological landscape.

Artificial Intelligence (AI) stands poised to revolutionize countless aspects of our lives, offering solutions to complex challenges and unlocking unprecedented possibilities. However, its rapid development and deployment present us with a critical juncture: how do we ensure AI serves humanity ethically, responsibly, and sustainably? This whitepaper delves into the multifaceted landscape of AI, exploring its transformative potential while acknowledging the inherent risks and challenges it poses.

We are now at the pivotal stage with AI technology adoption. There exists today a race amongst nations and organizations to develop technology that would be fundamental to AI and finding new applications for AI technology monetization. But as with any general-purpose technology adoption in the past, AI brings its own set of challenges with it. Since AI systems train over user data and make decisions based on machine learning, there are challenges associated with AI governance and ethical decision making. Thus, allowing for AI innovation to happen without any harmful effects necessitates a good Responsible AI framework. Sufficient guard rails that involve a mix of technology, standards, and regulations need to be put in place as safeguard measures. This report offers a high-level perspective on the various aspects of Responsible AI.

From navigating geopolitical disparities in AI development to addressing concerns about job displacement and data privacy, we aim to shed light on the complex issues that demand attention. By fostering a comprehensive understanding of these challenges, we hope to inspire collaborative action and guide the development of responsible AI framework; with sufficient guard rails that involve a mix of technology, standards, and regulations for evolution of AI for the benefit of all.

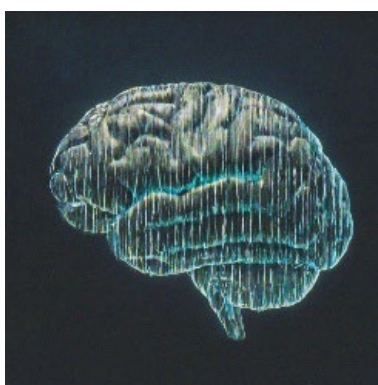
Disclaimer

5GIF extends our heartfelt gratitude to Qualcomm in driving the development of this whitepaper on AI and appreciates other 5GIF members for their feedback and suggestions to make the paper valuable. Their collective expertise and leadership have been instrumental in shaping the insights presented herein.

Introduction

Artificial Intelligence (AI) is a branch of computer science focused on creating systems that can perform tasks typically requiring human intelligence. If you've used Alexa, Siri, navigation maps, chatbots, or translation tools, you've encountered AI. AI aims to enhance human problem-solving abilities, but it's drawing regulatory attention because these tools are becoming more intuitive and autonomous.

AI is no longer just a futuristic idea; it's here and is changing our world now. From personalized suggestions to medical breakthroughs, AI is transforming industries and daily life. However, with this power comes great responsibility. As AI systems become more advanced and capable of making decisions that affect people, it's crucial to develop and use them ethically. This paper explores the balance between AI innovation and responsibility, looking at real-world uses and potential risks to guide AI development for the benefit of everyone.



AI - the next leap in Human-Machine interaction

AI refers to a computer system's ability to perform tasks that traditionally required human intelligence. It involves developing programming systems that analyse data and follow rules to solve complex problems, make smart decisions and continually learn from experiences. This enables computing machines to simulate collective human intelligence and problem-solving capabilities, which when automated has the potential to revolutionize multiple industries.

Artificial Intelligence (AI) is no longer science fiction; it's rapidly reshaping our world. From self-driving cars navigating busy streets to virtual assistants anticipating our needs, AI is transforming industries and permeating daily life. This evolution brings us closer to a future where machines understand us better, learn from us, and collaborate with us in unprecedented ways. A future world, AI-powered healthcare systems personalize treatments, where educational platforms adapt to individual learning styles, and where smart cities optimize resources for a more sustainable future. These are just glimpses of the transformative potential of AI.

Not just interacting with machines, creating new content using *Generative AI*, to generate new content like text, images, music, audio, and videos that trained over large volume of existing public information like blogs, discussion forums, websites, social media, technical papers and books of various genre.



AI encompasses several subfields and technologies, such as machine learning (ML), natural language processing (NLP) and Computer Vision (CV).

AI: Unveiling its Multifaceted Nature: While we often use "AI" as a single term, it encompasses a diverse range of technologies working together. Think of these subfields as specialized tools within the broader AI toolbox:

- **Machine Learning (ML):** At its core, ML empowers computers to learn from data without explicit programming. Imagine feeding a computer thousands of images labeled "cat" and "dog." Over time, the ML algorithm can identify patterns and features, eventually learning to distinguish between the two on its own. This ability to adapt and improve through experience is what makes ML such a powerful force in AI. Techniques like Deep Learning take ML one step further by mimicking the neural networks of the human brain and its learning process. This allows machines to process large sets of unstructured data by making non-linear, complex correlations by employing feedback loops.
- **Natural Language Processing (NLP):** This subfield focuses on enabling computers to understand, interpret, and generate human language. From chatbots that converse with us naturally to translation tools breaking down language barriers, NLP bridges the gap between human communication and machine comprehension. NLP also involves a wide range of technological components like speech recognition, natural language representation of computer programs too, convert instructions in natural language into intents and actions to perform functions in machine or subsystem.
- **Computer Vision (CV):** This area equips computers with the ability to "see" and interpret visual information. Just like humans use their eyes to understand the world around them, CV algorithms analyze images and videos, recognizing objects, scenes, and even emotions. This technology is driving innovations in fields like self-driving cars, medical imaging analysis, and security systems.

AI in Telecom and Future Networks

AI will play a pivotal role in shaping the future of 5G and 6G networks. The growing complexity of network infrastructure, densification of networks, and tens of billions of connected devices make it impractical and beyond human capacity to manage and operate networks using only traditional network management methods and static policies. AI and automation are helping networks overcome this immense challenge¹. By intelligently optimizing network performance, resource allocation, and traffic management, AI will ensure seamless connectivity and minimize latency. Additionally, AI-powered systems will enhance network security by detecting and mitigating cyber threats, safeguarding sensitive data and network integrity. Furthermore, AI can tailor network services to individual user preferences, delivering personalized content and experiences. Moreover, AI can automate routine network tasks, reducing operational costs and improving reliability. By processing vast amounts of data in real-time, AI can gain valuable insights into network behavior, enabling proactive maintenance and optimization.

The integration of AI in telecom is not just about improving current operations but also about paving the way for future innovations and services. As networks evolve with technologies like IoT and edge computing, AI will be essential in managing the increased complexity and delivering seamless, high-quality experiences to users.

AI systems employ algorithms that review, and model based on the copious quantities of data to identify and follow patterns using these techniques and then make data-driven decisions. Understanding these subfields is crucial as we explore the ethical implications and potential impact of AI on our lives.



Challenges associated with the Use of AI

AI holds immense promise for transforming countless industries and aspects of our lives. From revolutionizing healthcare to accelerating scientific discovery, its potential benefits are vast and far-reaching. However, navigating this rapidly evolving landscape presents significant challenges that must be addressed proactively. While AI has been known to exist for more than half a century, it is finding real life applications in the last half

¹ <https://www.ericsson.com/en/ai>

decade primarily due to large private investments. Surprisingly, these investments are known to be confined in a few select countries. Now just imagine if the first movers find applications of AI in fields like nuclear physics, space studies, molecular medicine, or another critical field of relevance to mankind's existence. The uneven distribution of resources and expertise, with a few countries leading in AI research and development, while others lag due to limited investments and a shortage of skilled professionals. This disparity threatens to widen the gap between technological haves and have-nots, potentially creating new geopolitical tensions. AI as a technology itself can singularly define the next superpower. While open-source AI; that includes datasets, models, ready-to-use interface and models to be used in public and local compute platforms are just catching-up², their adoptions are still catching up.

Among other concerns, establishing clear regulatory frameworks for AI is proving difficult. Governments worldwide are grappling to keep pace with advancements, leading to a fragmented approach to regulation. This lack of unified policy creates uncertainty for developers and raises concerns about potential misuse.

Furthermore, the automation potential of AI raises concerns about job displacement. Many fear that widespread adoption of AI could lead to significant job losses across various sectors, particularly those involving repetitive tasks.

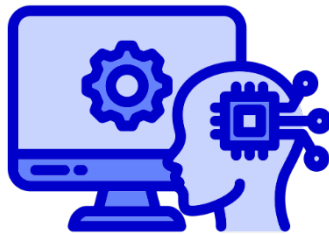
Countries with large population that rely on certain blue-collar jobs need to invest on reskilling to stay relevant and adopt to the new norm. This necessitates proactive measures to reskill and upskill, ensuring that the workforce gets adapted and contribute to the AI enabled economy.

Other challenges of AI are with the management of data in terms of security and privacy. The use of AI in applications like surveillance and facial recognition technologies raises ethical concerns about potential misuse and infringement on individual privacy rights. Striking a balance between leveraging AI's capabilities and safeguarding against these risks requires robust security measures, transparent governance frameworks, and ongoing public discourse.

There are risks leading to harm for individuals (privacy, security, financial, etc.) and to society. As AI algorithms process vast amounts of sensitive data, ensuring its confidentiality, integrity, and availability becomes paramount. Malicious actors could exploit vulnerabilities in AI systems to gain unauthorized access to data, manipulate algorithms for nefarious purposes, or launch sophisticated cyberattacks. Depending on who one talks to, the risks associated with the widespread adoption of AI can be broadly categorized into one of the below:

² <https://huggingface.co/>

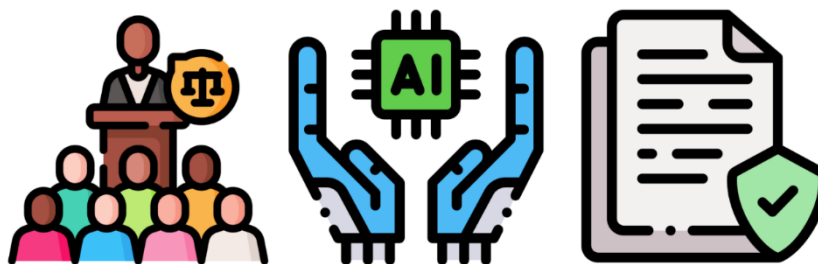
- i. **AI being used for malicious purposes:** consider the case of AI tools employed to learn a personality (say, politician) and mimic their behavior (in a bad manner) or creating deep fake video content of popular personalities (e.g., actors, CXO's of companies) with an intent to tarnish their reputation,
- ii. **Organizational challenges of AI:** Implementing AI in organizations requires a significant investment of time, resources, and effort. Any effort to cut corners will lead to sub-optimal designs. Issues such as bias in algorithms, data privacy concerns, and transparency in AI decision-making processes pose significant challenges. Standards defining the *right approaches* will be needed for addressing these challenges,
- iii. **Rogue AI elements:** AI algorithms are designed to learn and mimic human intelligence. At the end of the day, this is human code written to mimic human behavior (both prone to mistakes) which can result in producing unintended consequences.



We cannot predict how organizations, or for the same reasons certain administrations will react or respond to the adoption of AI in various sectors or even AI based applications. Since AI performance improves with learning, its implications may not be evident until the AI systems get deployed. Today AI tools are finding applications in many economic activities including online shopping and advertising, web search, digital personal assistants, language translation, smart homes and infrastructure, health, transport, and manufacturing. We cannot predict how organizations, or for the same reasons certain administrations will react or respond to the adoption of AI in various sectors or even AI based applications. Since a machine that lack's empathy was trained to learn from human behavior and then made to make decisions that will impact humans, there is consensus today that the actions/intents leading to such decisions needs to be regulated in ensuring that AI systems benefits humanity and in preventing any potentially damaging impacts. Responsible AI, which corresponds to the approach of developing and deploying AI systems to function in a safe, trustworthy, and ethical manner was in the developing. The rest of the article focuses on developing Responsible AI concepts.

The guiding principles of AI ethics³ include the following:

- Fairness: Datasets used for training the AI system must be given careful consideration to avoid discrimination,
- Transparency: AI systems should be designed and deployed in an explainable manner (allow understand how the algorithms work),
- Non-maleficence: AI systems should be designed in a manner to avoid harming individuals, society, or the environment,
- Responsibility: Developers, organizations, and policymakers must ensure AI is developed and used responsibly,
- Privacy: AI must protect people's personal data, which involves developing mechanisms for individuals to control how their data is collected and used,
- Robustness: AI systems should be resilient to errors, adversarial attacks and unexpected inputs.
- Inclusiveness: AI systems should be designed and deployed to accommodate diverse perspectives, needs and experiences, while being made available to everyone in society.



AI-Ethics Watchdogs

AI-Ethics⁴ watchdogs could be organizations, individuals, or entities that monitor and evaluate the ethical practices and behaviors of businesses, governments, and other institutions. Some firms may have or are likely to build internal watchdog teams that may monitor the use and development of AI within their organization. The role of such watchdogs should align with established ethical standards and societal expectations.

³ <https://www.iso.org/artificial-intelligence/responsible-ai-ethics>

⁴ <https://www.envisioning.io/vocab/ai-watchdog>

Guidelines for a Responsible AI System development

Summarizing on the previous sections, it is evident that the ethical and responsible development of AI demands that guidelines on the governance framework, safety elements and audit mechanism are needed⁵.

- *AI Governance*, refers to a collection of legal frameworks, policies, and best practices that serve as guardrails to ensure that AI technologies are researched, developed, and used in a way that minimizes potential risks and maximizes intended benefits,
- *AI Safety*, refers to developing technologies and governance interventions to prevent harms caused by AI systems that are aligned with human values, transparent, and accountable,
- *AI Assurance* refers to means by which organizations can then measure whether systems are trustworthy and demonstrate this to governments, regulators, and the market.

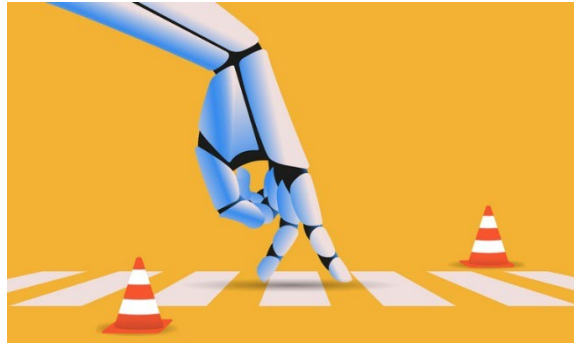
AI Governance

As AI tools become more powerful and prevalent, it will find widespread applications in sectors like healthcare, education, law, transportation, and finance. There are several challenges associated with such adoption:

- If trained using unscrutinized data, AI can replicate harmful biases about race, religion, upbringing, or other human characteristics. This could be potentially disastrous if embedded in human-centered applications like healthcare, recruitment, and law. It is therefore important for AI systems to adhere to certain rules,
- If not trained over an exhaustive set of data and used to make decisions in a critical application such as healthcare, a wrong decision can lead to life-threatening consequences. It is therefore important for AI systems to be transparent and accountable for it to be safe and reliable,
- A key ethical concern surrounding AI is privacy. With AI systems collecting vast amounts of data from databases worldwide, there is a need to ensure that personal information is protected and used responsibly. For example, facial recognition technology, often used in security systems or social media platforms, raises questions about consent and potential misuse.

The importance of ensuring the quality and reliability of AI systems cannot be overstated. For efficient governance, there need to be guardrails that need to be put in place to prevent AI (as a technology) from creating harm. This is essential for maintaining public trust in AI, and to ensure that AI-enabled technology operates safely within ethical and legal boundaries.

⁵ <https://www.nist.gov/artificial-intelligence>



AI guardrails can be implemented by employing a combination of technical controls, safety standards, policies, and laws. Technical controls are embedded during the AI system design and development phase itself and need to comply with safety standards in place (e.g., EU AI Act⁶). In contrast, policies are internal or external operational guidelines, and laws are enforceable regulations enacted by governments.

- Technical controls need to comply with safety standards and include embedding technical controls directly into AI workflows as operational processes that are integral to the AI system functioning. These controls include but should not be limited to,
 - Tagging AI generated content (e.g., watermarking),
 - Training, validation, and testing of AI system designs,
 - Incorporating rule-based safeguards to AI systems,
 - Adding security guidelines and protocols to AI systems,
 - Keeping a robust incidence reporting mechanism.
- Including organization specific policy control mechanisms to determine how AI workflows are designed and managed. These include guidelines for,
 - Data governance,
 - AI fairness and accountability,
 - Data security,
 - Adherence to industry regulations,
 - Intellectual property protection on AI generated content,
 - The safety of AI applications.
- AI regulations need to be enacted, and formal standards developed to assess compliance with prevailing laws and regulations.

These guardrails are complimentary in nature, and when combined can help ensure the responsible AI development and deployment. Complying with data governance rules and privacy regulations as well as prioritizing safety, trustworthiness and transparency are also important to the future of AI Governance.

⁶ <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

Note: Sometimes AI Governance is confused with Data Governance. AI Governance are used to ensure that AI models are deployed, operated, and monitored ethically, transparently, and in line with societal values and organizational priorities, whereas Data Governance deals with the managing and regulating the use of data within organizations and ensures that the data feeding into AI models is accurate, secure, and trustworthy. Data governance covers the full data lifecycle, including AI, and is a fundamental pillar of AI governance.

AI Safety

AI safety is an interdisciplinary field focused on preventing accidents, misuse, or other harmful consequences arising from AI systems. AI Safety gained prominence due to the rapid progress and growing public concerns on the potential dangers associated with Gen AI. It encompasses several key aspects that include,

- Machine Ethics and AI Alignment: these aim to ensure that AI systems are moral, beneficial, and aligned with human values (e.g., rule based ethical guidelines),
- Risk Monitoring and Reliability Enhancement: monitoring AI systems for risks and enhancing their reliability (e.g., detecting biases, ensuring robustness, and minimizing unintended consequences). It also involves monitoring to ensure that AI systems remain aligned to objectives and are controllable,
- Norms and Policies: developing norms and policies that promote safety in AI development and deployment.



AI safety research involves efforts to prevent critical system failures, addressing bias, and safeguarding against misuse. Some topics of research on AI safety include,

- i. Data Poisoning: developing methods to protect AI models from maliciously manipulated training data (safeguarding from biased or unsafe behavior),
- ii. Algorithmic Fairness: ensuring that AI systems do not discriminate against specific groups based on race, gender, or other protected characteristics,
- iii. Explainable ML: making AI models more interpretable (understand their decisions and identify potential biases),
- iv. Gender Bias: detecting and mitigating biases related to gender,

- v. Curating training data: detecting when AI models encounter out-of-distribution training data (prevent unexpected and unsafe behavior).

In addition, AI safety practices need to be incorporated into organizational workflows for responsible AI development. A comprehensive approach that combines technology, expertise, and process improvement ensures effective AI safety practices. Here are some best practices:

- Identify a clear use case, and start training the AI system using curated data,
- Leveraging AI tools using continuous Human-Machine interactions,
- Automate repetitive tasks related to safety monitoring, risk assessment, and compliance to improve efficiency and accuracy,
- Implement AI-powered predictive analytic systems that analyze real-time data and identify risks. This proactive approach helps prevent safety issues,
- Upskilling on AI in a regular cadence to keep updated on AI safety elements.

AI Assurance

AI Assurance corresponds to *audit* processes to be followed in declaring that the AI system conforms to predetermined standards, best practices, and regulations. The objective is to ensure optimal and risk-free AI systems. Key elements of assurance include the creation of overarching and sector-specific regulation, standards, and principles. This includes practical capabilities to extract, timely and understandable insights about underlying AI technologies.

Standardization in AI

Do note that there are several ongoing worldwide efforts to regulate AI (e.g., EU AI Act)⁷. Currently governments are playing catch-up as AI applications are developed and rolled out. Despite the transnational nature of this technology, there is no unified policy approach to AI regulation, which indeed is a point of bother. Current trends are aimed at promoting the responsible growth of technology and are therefore limited to mere legal compliance towards ensuring AI's social responsibility, targeting to safeguarding against financial, legal, and reputational damage. Interim steps including mandatory guardrails for high-risk AI systems, best practices on the use of AI, and a voluntary labelling and watermarking scheme for AI-generated materials are also followed by some companies.

⁷ <https://www.aiprm.com/ai-laws-around-the-world/>

Another area of AI standards focus is AI Benchmarking. AI-focused benchmarks are employed by organizations to compare their specific implementations. These include standardized metrics to evaluate and compare the performance of various algorithms, model architectures, and hardware run platforms. Note that these benchmarks typically serve commercial interests, and therefore serve more akin to a beauty contest.

The future of AI governance depends on collaboration among governments, organizations, and stakeholders in developing the necessary techno-regulatory framework that protects the public while fostering innovation. Creating AI guardrails that everyone agrees with is going to be incredibly challenging. This is where standards can help. International Standards, that safeguard and propel the principled application of AI technology are the need of the hour.

6G networks are expected to natively implement artificial intelligence (AI)/machine learning (ML) functionalities demonstrating meaningful performance gains against competitive benchmarks. These functionalities will be distributed across the protocol stack and will likely include at least some of those studied and specified in the latest 3GPP releases of 5G Advanced, which has paved the way for 6G.

AI standards will eventually define the mandatory and optional requirements, best practices, and constraints that need to be satisfied for an AI method, algorithm, application, or system. Such standards shall cover technology elements, data governance, and other areas including privacy, bias, transparency, and accountability. Adherence to formal standards will help to bridge the gaps in regulation, allowing policymakers to establish consistent and auditable data and processes. It is understood that sector specific regulations will be needed, eventually, which would be tailored to specific disruptions with AI.

Artificial Intelligence will impact on almost every aspect of future ICT systems. Accordingly, many standards-developing organizations around the globe are adopting a distributed approach to AI-related standardization activities. Currently, numerous standards organizations worldwide are working on norms for AI technologies and processes.

1. At the International level:

- **ISO/IEC JTC 1/SC 42- Artificial intelligence**⁸: This Joint technical committee is responsible for standardization in the area of AI. To date, it has published over 30 standards, covering various aspects of AI technologies, processes, and applications.

⁸ <https://www.iso.org/committee/6794475.html>

2. In India:

- **BIS LITD30 on Artificial Intelligence:** LITD 30 on AI within BIS is responsible for standardization in the area of Artificial Intelligence and Big Data. It serves as India's national mirror committee for ISO/IEC JTC 1/SC 42.
- **TEC Convergence & Broadcasting (C&B) Division**⁹: Telecommunication Engineering Centre (TEC), Department of Telecommunication (DoT) is also engaged in AI related work through its C&B Division. It has recently published TEC 57050:2023 on “Fairness Assessment and Rating of Artificial Intelligence Systems”.

3. In Europe:

- **CEN-CENELEC JTC 21 on Artificial Intelligence**¹⁰: This JTC produces standardization deliverables in the field of AI and related use of data, as well as provide guidance to other technical committees concerned with Artificial Intelligence. The JTC also considers the adoption of relevant international standards and standards from other relevant organizations, like ISO/IEC JTC 1 and its subcommittees, such as SC 42 Artificial intelligence.
- **ETSI Technical Committee on Securing Artificial Intelligence (ISG SAI)**: ETSI TC SAI has a key role to play in improving the security of AI through production of high-quality technical standards; the TC SAI will create standards to preserve and improve the security of new AI technologies. The ETSI ISG SAI focuses on four key areas: using AI to enhance security, mitigating against attacks that leverage AI, securing AI itself from attack and Societal security and safety aspects of the use and application of AI.
- **ETSI ISG on Experiential Networked Intelligence (ENI)**¹¹: ETSI ISG ENI defines a Cognitive Network Management architecture. This is using Artificial Intelligence (AI) techniques and context-aware policies to adjust offered services based on changes in user needs, environmental conditions and business goals.
- Other Technical Committees (TC) and Industry Specification Groups within ETSI which are currently working on AI-related topics include ARF (Augmented Reality Framework)¹², CIM (Context Information Management)¹³, CYBER (CyberSecurity)¹⁴, eHealth¹⁵, INT (Core Network and Interoperability Testing)¹⁶, MEC (Multi-access Edge Computing)¹⁷, NFV (Network Functions Virtualization)¹⁸, PDL (Permissioned

⁹ <https://tec.gov.in/ai-fairness>

¹⁰ <https://www.cencenelec.eu/areas-of-work/cen-cenelec-topics/artificial-intelligence/>

¹¹ <https://www.etsi.org/technologies/experiential-networked-intelligence>

¹² <https://portal.etsi.org/tb.aspx?tbid=858&SubTB=858>

¹³ <https://portal.etsi.org/tb.aspx?tbid=854&SubTB=854>

¹⁴ <https://portal.etsi.org/tb.aspx?tbid=824&SubTB=824,856>

¹⁵ <http://portal.etsi.org/tb.aspx?tbid=696&SubTB=696>

¹⁶ <https://portal.etsi.org/tb.aspx?tbid=715>

¹⁷ <https://portal.etsi.org/tb.aspx?tbid=826&SubTB=826,874#/>

¹⁸ <https://portal.etsi.org/tb.aspx?tbid=789>

Distributed Ledger)¹⁹, SmartM2M (Smart Machine-to-Machine)²⁰, ZSM (Zero-touch Service Management)²¹, etc.

Conclusion

AI has made remarkable progress in recent years, transforming numerous fields. While AI holds immense promise, it's crucial to implement thoughtful risk management strategies to ensure its safe and ethical application (consider for e.g., the rise of Generative AI). Responsible AI encompasses several key aspects: developing technologies with ethical considerations in mind, adhering to industry standards and complying with evolving regulations. AI regulation involves public sector policies and laws designed to promote and oversee the development and deployment of AI systems. These broader efforts aim to regulate and ensure that AI technologies are used ethically, transparently, and in a way that respects fundamental rights individual safety.

This field is constantly evolving, with policymakers continuing to refine their approaches as technology advances. Remember that responsible AI is an ongoing commitment. Organizations must continuously adapt and improve their practices to align with the latest evolving standards and regulations.

-- ## --

¹⁹ <https://portal.etsi.org/tb.aspx?tbid=873&SubTB=873#/>

²⁰ <https://portal.etsi.org/tb.aspx?tbid=726&SubTB=726>

²¹ <https://portal.etsi.org/tb.aspx?tbid=862&SubTB=862,863>