STANDALONE REPORT

GLOBAL STANDARDS MAPPING INITIATIVE 4.0
NOVEMBER 2023

ARTIFICIAL INTELLIGENCE (AI) & CONVERGENCE

GLOBAL BLOCKCHAIN BUSINESS COUNCIL

DC Location:
1629 K St. NW, Suite 300
Washington, DC 20006

Geneva Location:
Rue de Lyon 42B
1203 Geneva
Switzerland

GSMI 4.0 IN-DEPTH REPORT

ARTIFICIAL INTELLIGENCE (AI) & CONVERGENCE

AI & CONVERGENCE OVERVIEW

What is AI?

Artificial Intelligence (AI) refers to the use of technology to simulate human cognitive functions, enabling computers and machines to perform tasks such as problem-solving, decision-making, understanding and producing natural language, recognizing patterns, and adapting to changing environments.

Human-machine interactions can be direct, where humans engage with AI interfaces, or indirect, where AI systems work behind the scenes to enhance productivity or decision-making.

AI can take advantage of and create synergies with other new technologies. For example, blockchain can record data, which may someday be utilized by AI to draw patterns and help make informed decisions based on validated data. AI is also used for monitoring transactions in both decentralized finance and high frequency trading in traditional financial products. Cryptocurrencies can be used to pay for AI processing power. Internet-of-things tools and sensors can provide vast amounts of data that are necessary for AI training and processing. Cloud technology, with its vast processing power, is used by many AI models and applications.

This working group will discuss the major facets of the AI landscape today, reviewing the main considerations for companies and organizations seeking to deploy AI innovations to take into account, discussing a select number of use cases for AI today, and the policy implications of such uses.

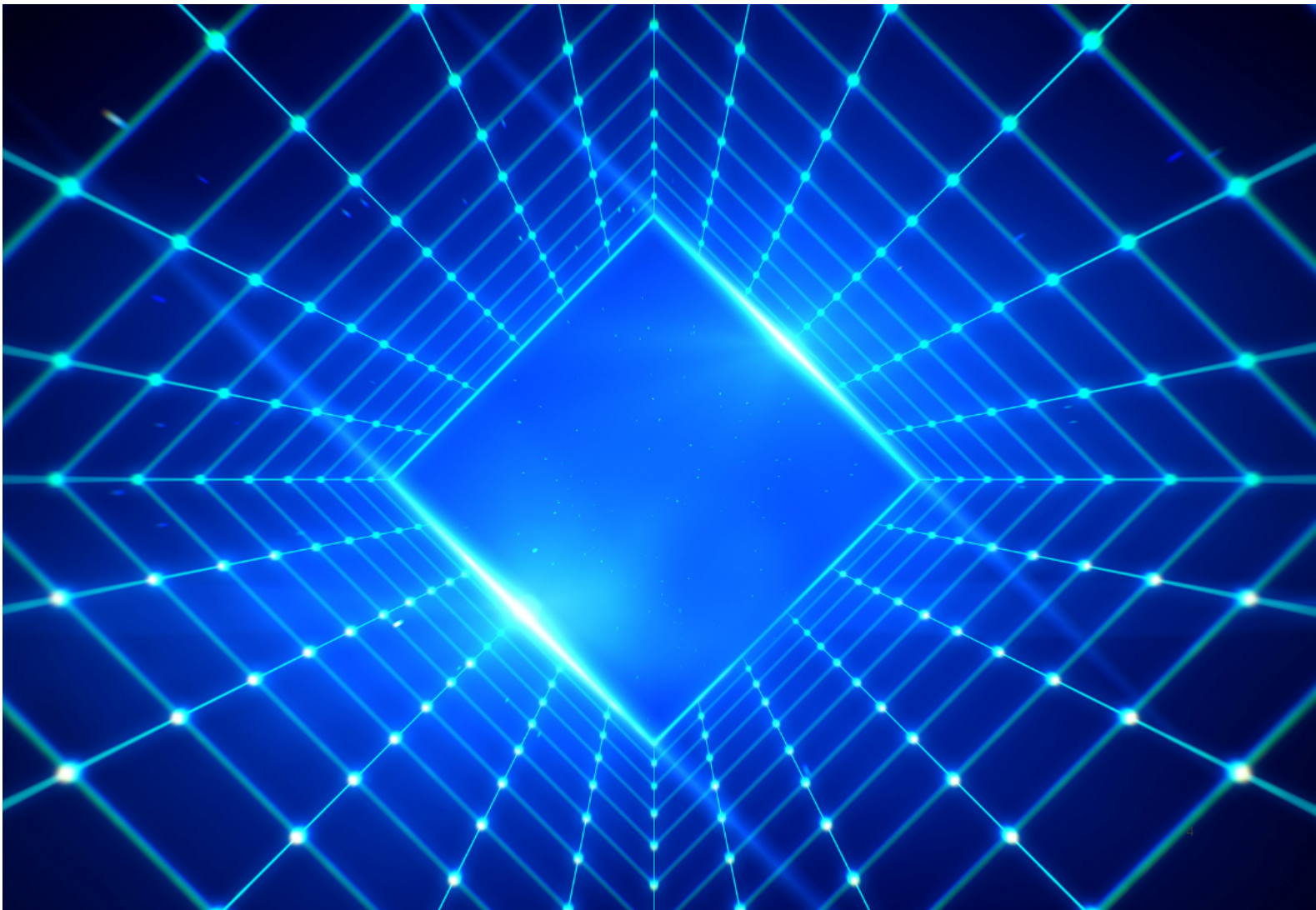
Opportunities for AI

With the rapid expansion of AI across industries, AI has become a disruptive force that is reshaping how businesses operate and how individuals interact with machines. This new field is full of possibilities, such as:

- **Automation -** AI can automate repetitive tasks, improving efficiency and reducing human error.
- **Data Analysis -** AI can analyze vast datasets quickly, enabling data-driven decision-making in various domains.

- **Personalization -** AI can tailor user experiences, such as content recommendations or product suggestions, based on individual preferences and behaviors.
- **Innovation -** AI can drive innovation by enabling the development of new products, services, and solutions that were previously unattainable, through automation of testing or generation of new formulae.
- **Improved Decision-Making -** AI can provide insights and predictions to assist human decision-makers in various fields, from healthcare diagnostics to financial forecasting.
- **Enhanced Safety -** In industries like transportation, AI-powered systems can enhance safety through features like autonomous driving and predictive maintenance.
- **Content Creation -** AI algorithms can be utilized to create content in literature, art, and a range of human activities. For instance, Creative Commons uses generative AI for content creation, in ways that support its mission to support open access to education and creative works, which can be shared and built upon legally based on its licensing.[1]

AI is ultimately an enabler of technological solutions in human-centered and social contexts. It enables a wide array of human-machine interactions with multiple possibilities, roles, and functions. AI can help address the most complex problems facing humanity, with adequately defined parameters, dimensions, and values to include in the algorithms on which it runs.



Risks of AI

It is important to remember that AI systems do not actually "reason." in the same way that human beings do, which requires emotional associations and logical cognition. AI models merely "think" in the sense that they perform processes that require decision making based on datasets and information, and thus they mimic a narrow part of human thought processes. For instance, both AI software and human "thinking" can come up with a word that rhymes with another word based on an existing dataset of vocabulary from which to choose similar sounding words.

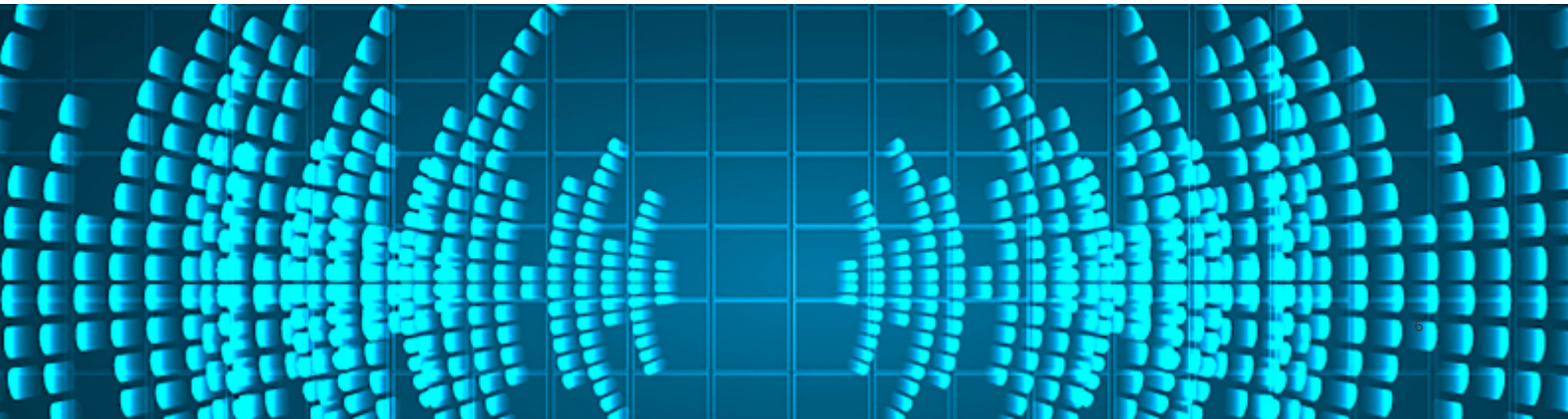
> On the other hand, AI does not have and cannot mimic human reasoning, as logical cognition and emotion are inherently human qualities. When asked why blue is a beautiful color, while humans may answer with an emotional association of the color to a bright blue sky, or a feeling or mood, an AI model would answer based on its past dataset.

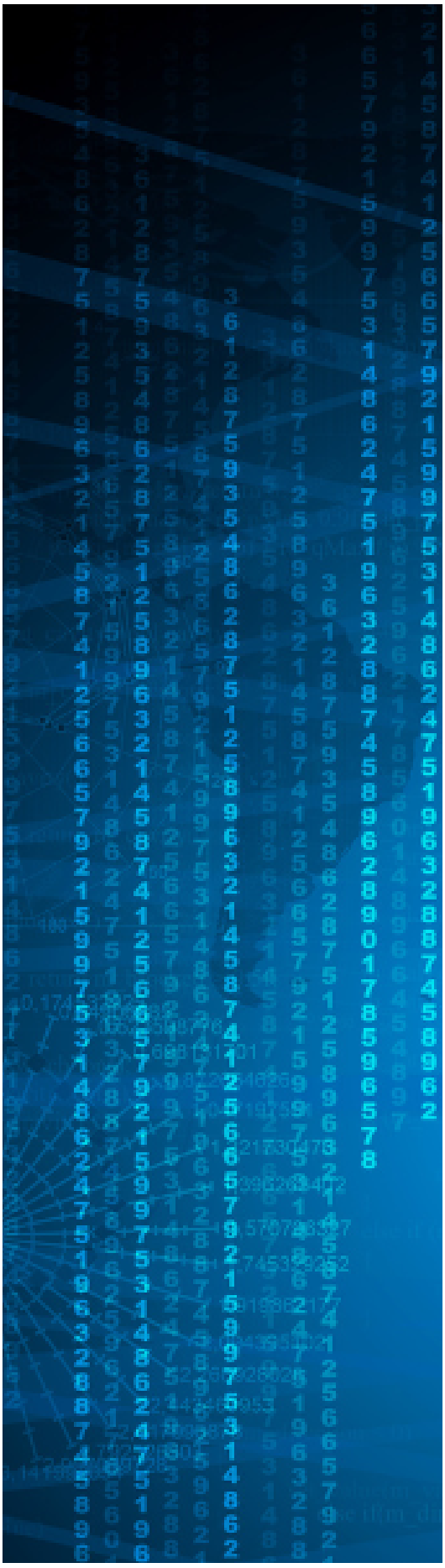
It is also humans that program the information and datasets that go into AI systems, which can contain emotionally-driven biases or other unwanted implications, which are then perpetuated as AI uses predictive models to "think" of which information from the given dataset is relevant to determine a given action. AI models are predictive, taking in vast amounts of data to recognize patterns and predict the next sequence in the pattern, which data elements are anomalies, or what should be the result from a series of prompts. An AI model makes predictions based on the data that is used to train it, and eventually it learns from its interactions with human users. Therefore, its predictions will not have any inherent moral code or guidelines unless those are built into the system.

Moreover, AI models may experience hallucinations, where responses generated based on data inputs produce false or misleading information presented as facts. AI models, such as language models or image generators, may make predictions or produce outputs that incorrectly extrapolate on their training data, or are not fully grounded on their data inputs on which they rely.

AI therefore can have unintended consequences when unchecked. The complexity and lack of human oversight of AI systems can perpetuate unwanted behaviors and influence decisions with negative consequences for individuals, businesses, and human civilization. AI presents a series of risks such as:

- **Lack of Transparency -** AI systems and models can be complex and therefore hard to interpret, which makes decision-making processes and their underlying logic opaque.
- **Concentration of Power, Inequality, and Bias -** When AI relies on biased training data or algorithms, it can amplify and perpetuate societal biases. AI developments can disproportionately benefit wealthy individuals and corporations, aggravating the wealth gap and opportunities for social mobility. Concentration of processing power and data can not only increase social and economic inequalities when AI developments are dominated by a small number of large entities (e.g., corporations and governments), but also can have geopolitical effects. Concentration of models could have a wide range of unintended results, such as herding, and will also create single points of failure. In addition, concentration of market power could lead to monopolistic practices.
- **Lack of Privacy -** AI often collects and analyzes significant amounts of personal data, which needs to remain private and secure.
- **Ethical Issues -** Moral and ethical values embedded into AI systems can present significant ethical dilemmas, particularly in the context of decision-making processes with major consequences for people's lives.
- **Lack of Security -** Increasingly sophisticated AI models also raise security risks, including potential misuse. Hackers and bad actors can make use of AI for cyberattacks, bypassing security measures, and exploiting system vulnerabilities. Moreover, AI-driven autonomous weapons also can come into the hands of rouge nations and non-state entities, which further raise concerns given the potential loss of control by humans in critical decision making. A resulting AI arms race can promote rapid and unchecked development of AI with harmful consequences.
- **Concentration of Power -** Concentration of processing power and data can not only increase social and economic inequalities when AI developments are dominated by a small number of large entities (e.g., corporations and governments), but also can have geopolitical effects.
- **AI Dependence and loss of human connection -** If society becomes over reliant on AI systems, it can lead to a loss of creative initiatives, critical thinking, and human intuition, which are not only key to preserving human cognitive abilities but also for addressing the most pressing issues and human flourishing. Moreover, dependence on AI-driven communications and interactions could hinder levels of empathy, social skills, and human connection.





- **Job Displacement -** Automation driven by AI across industries can weaken the power of human workers and lead to job losses, especially affecting low-skilled workers.
- **Legal and regulatory challenges -** New legal frameworks and regulations are fundamental to address the novel issues posed by AI developments, such as liability and intellectual property rights, while protecting the rights of all citizens.
- **Manipulation and Misinformation -** The spread of AI-generated false content, such as deepfakes, can manipulate public opinion. Disinformation can threaten democracy and promote authoritarian political leadership (e.g., facist currents, ultranationalist ideologies, oppressive centralized autocracies, etc.) by influencing public discourse, spreading fake news, and undermining social trust. Extremist groups, criminals, and rogue states can manipulate groups of people for economic and political interests.
- **Existential Threats -** An artificial general intelligence (AGI) that surpasses human intelligence on various functions can present long-term concerns including a threat to our very existence, especially because AI may not be aligned with human interests, priorities, and values.
- **Super-Dominant AI Platforms -** A potential danger arises if AI development results in the existence of a few large-scale AI systems – such systems could become super-dominant, preventing competitors from gaining a foothold due to heavy investments in computing power, data and hardware that smaller companies are unable to duplicate. The creation of a few dominant systems could amplify other risks such as bias and act to further entrench companies that already have significant market share at the expense of potentially beneficial competitors.

In the worst-case scenario, the risks of unchecked AI can threaten to worsen wealth inequalities, weaken human agency, and even threaten human existence.

Blockchain Convergence and Opportunities

Blockchain technology can add a layer of trust for AI developments, which could draw patterns and guide informed decisions based on validated, immutable, and open data records.

- **Visibility on Algorithms -** The features of blockchain technology add a layer of transparency into AI developments, which can be fundamental to ensure safe and reliable outcomes, and ultimately safeguard trust. When people can comprehend the reasoning and processes with which an AI system arrives at conclusions and outcomes, greater trust and adoption can follow.

- **Transparency on Data Provenance -** Blockchain technology can provide transparency into the source of data sets, helping identify risks of biased, or narrowly focused data sources. A blockchain can document and validate relevant input data from an adequately identified and anonymized source, such as clinically and medically relevant patient data that an AI tool can rely on to identify cancer cells.

  Blockchain's promise in validating data provenance can address the inherent biases that can be present in AI filtering models. This can alert the need to make use of diverse training data sets and unbiased algorithms to ensure fairness in outcomes. Transparency, particularly with respect to data provenance, can also help identify false content in order, to preserve the integrity of information in our digital age.

- **Transparency on AI Outcomes -** Blockchain technology can also provide transparency into the uses and outcomes of AI, such that accurate records of AI-driven decisions and uses recorded on an open ledger can be preserved for monitoring and evaluation of results. For instance, both the use of an AI tool to show "yes/no" regarding whether medical imaging shows the presence of cancer, as well as the following decision to take measures accordingly, can be recorded accurately on a blockchain. The analysis based on the results can also be documented immutably and preserved.

  Therefore, blockchain technology can provide a transparent lifecycle of data sources, uses, and results. It can validate the provenance data going into AI tools, as well as the provenance of the output of AI tools, and track the success of outcomes based on AI-driven decisions. With better controls in place enabled by transparency regarding input data and outputs, it can be more feasible and realistic to determine the effectiveness of AI tools and reduce the risk of relying ineffective or under-performing tool.

- **Data Privacy and Security -** In addition to transparency, blockchain technology at the intersection of AI can also improve privacy protection mechanisms, which can safeguard the privacy of individuals while ensuring security and dependability of data. A range of privacy protecting techniques including data encryption, and fully-anonymized data sets, can greatly improve trust for functions such as authorization management, access control, data protection, network security, and scalability.[2]

- **Decentralization to address power concentrations -** Decentralization can prevent concentration of power and single points of failure, adding resilience and trust. Decentralized and collaborative developments toward AI are fundamental to avoid concentrations of power in AI that can further existing inequalities. On the other hand, decentralized AI developments can contribute toward inclusive relationships and economies.

- **New processes to preserve equality -** AI in convergence with blockchain technologies can also reinvent processes, such that automation doesn't lead to job losses where the same functions are replaced by machines. On the other hand, in decentralized economic interactions with new peer-to-peer possibilities and governance, more jobs can be created than those eliminated, while preserving inclusion and equality.

A BRIEF TAXONOMY

The field of AI is vast and encompasses various subfields, techniques, and applications. In that sense, a taxonomy is essential to provide a structured framework for organizing and categorizing these diverse elements. It helps researchers, educators, and policymakers to establish a common language and terminology for discussing AI concepts, methods, and technologies. This standardization enhances communication and understanding among individuals and organizations involved in AI research and development. This working group has also begun preparation of an AI taxonomy.[3]

Figure 1: Generative and Predictive AI are two alternative functions

Generative AI

Produces new content

Predictive AI

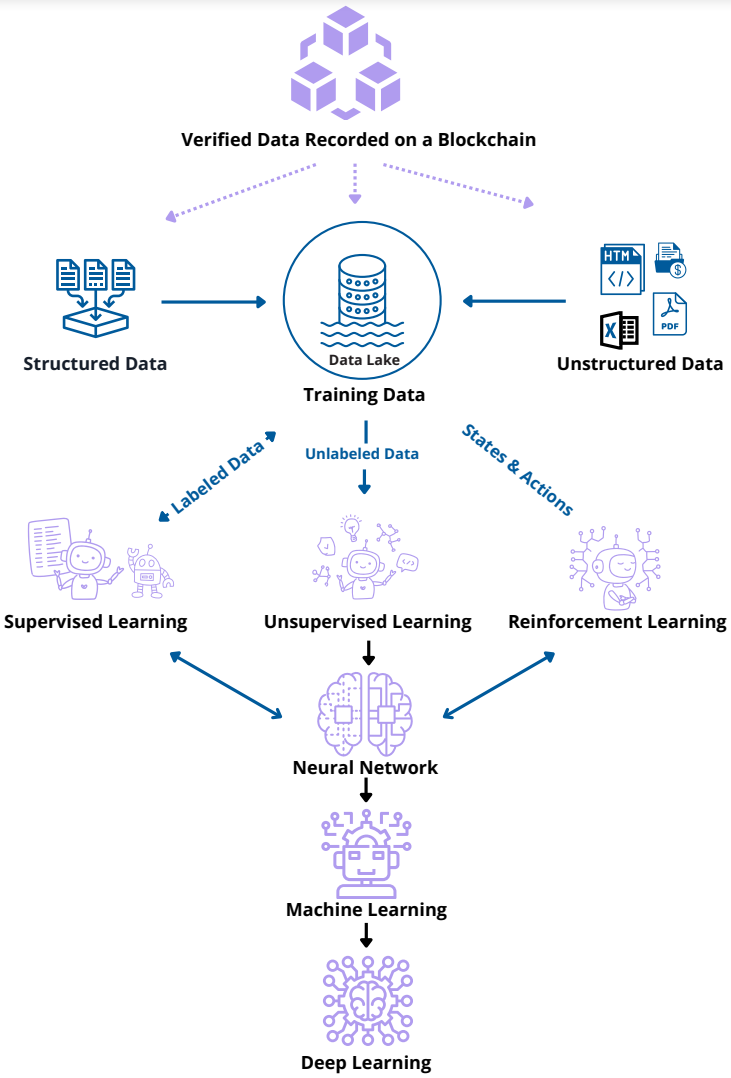
Anticipates future content based on past patterns

Much of AI perception today focuses on Generative AI, that is, models that are trained to generate new original content based on natural language input. Some Generative AI is in the form of chatbots, which generate responses to user input in a way that mimics predictive text, but in a much more powerful way. Other types of Generative AI allow users to describe a desired output in normal everyday language, and the model can respond by creating appropriate text, images, sounds, videos, or even code output. Yet there is much more to AI than just generative models.

What this paper refers to as "predictive AI" is a method of data analysis which recognizes patterns and analyzes those patterns to make predictions about future outcomes using historical data combined with statistical modeling, data mining techniques and machine learning. This allows for prediction of trends and risks, identification of anomalies, and categorizing of data.

Different use cases for AI call for different models, where those models may be generative, predictive or a combination of both.

Figure 2: How AI functionalities come together



AI functionalities demonstrated above can be seamlessly integrated with blockchain technology, which can serve as the immutable trusted record on which AI models source their data. Sourcing inputs from decentralized ledgers can also address the risks of power concentrations to better safeguard fairness, equality, and human freedom. Data originating from records on a blockchain can add a layer of auditability and validation for accuracy, and ultimately greater trust in the AI models.

USE CASES

When assessing new AI tools generally, companies and organizations must assess their own understanding of the technology and how it can be used to further business goals in compliance with law and with the organization's goals. Each new use will require an evaluation of the AI model in the context of the organization's overall operations and proposed use. This paper provides an overview of certain use cases and matters that should be considered when adopting AI tools for these uses as examples of how organizations should consider implementing such tools. However, there are certain considerations that will be applicable to almost all AI uses. General considerations for AI implementations include the following:

- **Regulatory Compliance -** Each organization implementing an AI tool should assess its regulatory posture with respect to how that tool will be used. The use cases listed below are sector-specific examples of how organizations operating in different industries might consider the use of a new AI tool. However, generally applicable regulations and policies such as anti-discrimination, sanctions, and data privacy will be relevant to all uses.

- **Protection of Intellectual Property -** Generative AI tools can create new intellectual property – but whether this property is protected under applicable laws, and the rights the user of the tool has to newly created property, should be considered. In addition, use of protected intellectual property as inputs or training data creates infringement risks.

- **Safety Considerations -** Depending on the specific applications, safety can be a paramount concern. For AI-powered systems, particularly those in sectors like autonomous vehicles, healthcare, or manufacturing, assessing and ensuring the safety of the technology is critical. This includes rigorous testing, validation, and verification processes to minimize the risk of accidents, injuries, or other adverse outcomes caused by AI failures. Other types of AI tools must also take safety into account – for example, chatbots, especially those targeted to vulnerable audiences, need to be assessed to ensure that they do not produce harmful content. AI tools should be subject to regular testing and audits to detect and mitigate biases, errors, and unintended consequences[4].

- **Data Privacy -** Organizations must assess how an AI tool will use and store data in order to ensure compliance with data protection regulations such as GDPR or HIPAA. In addition to restrictions on sharing and use of personal data, the GDPR grants its subjects the right to not be subject to decision based solely on automated processing which produces legal/significant effect, subject to certain exceptions.

- **Information Security -** Organizations should implement strong cybersecurity measures to protect AI and blockchain systems from attacks (robustness), including encryption, multi-factor authentication, and intrusion detection systems. Ensure that AI tools are secure and free from vulnerabilities that could be exploited by malicious actors.
- **Concentration Risks -** AI tools can also create risks associated with a lack of robustness, alignment and/or controllability of strong AI systems. The upcoming UK AI Safety conference will focus on this. So is the Global Partnership on AI (GPAI) mandated by the G7 to look at AI and in particular generative AI safety solutions (G7 Hiroshima AI Process).
- **Interoperability -** Evaluate how AI and blockchain solutions will integrate with existing systems and technologies within the organization. Compatibility and interoperability can be critical for successful implementation.
- **Data Governance:** Establish robust data governance practices to maintain data integrity, quality, and traceability throughout the AI and blockchain lifecycle.[5]
- **Ethical and Moral Considerations -** Organizations seeking to implement AI tools must consider the moral ambiguity and ethical questions that AI can pose. Although government regulatory frameworks are expected to develop and eventually guide AI implementations with these ethical and moral considerations at the core, these regulatory developments are not likely to be drafted in the time needed to ensure the right safeguards today. Therefore, in absence the of comprehensive regulatory frameworks for AI as of now, organizations should evaluate how launching AI tools can raise ethical and moral uncertainties. They should also adapt to these risks accordingly, or to the extent possible take steps to minimize the risks.

Major AI Implementations Today

The below use cases are illustrative examples of how companies and organizations might consider the implementation of a new AI tool.

FINANCIAL/FINTECH USE CASES

I) Anti-Money Laundering/Know Your Customer (AML/KYC)

The methods currently employed by financial institutions to achieve compliance with Anti-Money Laundering ("AML") and Know-Your-Customer ("KYC") requirements tend to involve human review of transactions and customer identification materials, which injects inefficiency and human error into this important but costly function. Predictive artificial intelligence can materially improve these manual processes by allowing for a rapid and efficient review of large data sets via automated processes that mitigate the inaccuracy of manual human review to achieve more accurate results at a fraction of the cost. The large sets of data reviewed by an artificial intelligence can then be stored centrally via blockchain technology, allowing for ease of access to the results of such review as well as the data underlying those results. Combined with advances in digital identity, use of artificial intelligence for AML and KYC applications promise new levels of efficiency and accuracy.

However, financial institutions employing artificial intelligence to automate review of transactions and customer identification should ensure that some level of human review remains in place with respect to both the results produced by such review and the explainability of the decisions made by any artificial intelligence involved in these processes. Financial institutions should also consider how the models they employ are being trained and deployed, and how any data inputs to the model, or outputs from the model, are shared outside of an organization to ensure compliance with data privacy and confidentiality obligations with respect to customer and transaction data processed by the model.

I.I) AML AND KYC REQUIREMENTS FOR FINANCIAL INSTITUTIONS

The **Currency and Foreign Transactions Reporting Act of 1970** (the "Bank Secrecy Act") details the AML requirements imposed on financial institutions. At its core, the Bank Secrecy Act requires that financial institutions maintain records of all cash purchases of negotiable instruments, file reports of all cash transactions in excess of USD$10,000 per day, report any suspicious transactions indicative of money laundering or other criminal activities, and maintain a security program of policies and procedures designed to ensure compliance with Bank Secrecy Act requirements.[6] FDIC-supervised financial institutions are subject to additional reporting requirements in connection with any known or suspected criminal activity in connection with transactions conducted through such institutions.[7]

The **Financial Crimes Enforcement Network** ("FinCEN") promulgated the Customer Due Diligence Requirement for Financial Institutions (the "CDD Rule"), effective as of July 2016,[8] which requires certain financial institutions to identify and verify the identity of their customers, a process that has come to be known as "Know-Your Customer" or "KYC."

At its core, compliance with the CDD Rule's KYC requirements imposes on qualifying institutions the obligation to identify and verify the identity of customers and of customers' beneficial owners, to understand and develop respective risk profiles, and to conduct ongoing monitoring of suspicious transactions and customer identities.[9]

The compliance programs implemented by financial institutions today largely comprise a high degree of manual individual review of large sets of customer information and volumes of transactions on a daily basis. The manual nature of these processes, and the amount of customer information and transaction data requiring review, has produced an inefficient system whereby substantial time and effort is devoted to providing KYC information to financial institutions, validating the KYC information provided by customers, and subjecting large numbers of transactions to several layers of review and escalation as appropriate.

Unsurprisingly, these compliance efforts result in large costs (both financially and temporally) borne by financial institutions and by customers who may lack the sophistication and resources to effectively understand and meet the information requests they receive. The process further opens the door to inaccuracies resulting from human error and the large volume of data available; false positives and false negatives are bound to occur, especially in light of the growing sophistication of money laundering techniques and the proliferation of blockchain-based transactions in cryptocurrencies.

I.II) HOW CAN AI HELP?

Given the core issues with AML and KYC compliance today, it should come as no surprise that **AI can make a substantive impact on the speed, efficiency and accuracy of AML and KYC reviews.** Machine learning models can be trained to allow for rapid review of large data sets, and to screen customer and transaction data against a broader and more comprehensive list of data points (e.g. sanctions list, media, internal data points, watchlists etc.) improving accuracy and reducing human bias in review (though consideration should be given to bias in machine learning models as well, as discussed below) of transactions.

Furthermore, **AI can assist in expediting customer onboarding** by extracting and validating structured data in an automated and efficient manner, and reliably comparing them against trusted data sources for validation, significantly reducing costs associated with one of the most labor-intensive aspects of KYC. Furthermore, the ability to store and access KYC using blockchain technology can provide an ever-growing secure and robust source of data that can be used across institutions to reduce redundancy in process and lower costs for financial institutions.

I.III) CONSIDERATIONS FOR FINANCIAL INSTITUTIONS SEEKING TO UTILIZE AI IN AML AND KYC COMPLIANCE FUNCTIONS

While AI can substantially improve the efficiency and efficacy of AML and KYC compliance processes, financial institutions seeking to automate such processes through the use of AI should implement robust policies and procedures designed to ensure careful consideration and monitoring of the incorporation of AI into such processes. Such programs should require a balanced approach by an institution that includes careful human review of both inputs to and outputs from an AI model at critical points in the AML and KYC compliance processes and routine reviews of random samples of data reviewed by an AI model to validate any recommendations made by the model with respect to such data. Firms should additionally require that any AI technology it uses allow for sufficient explainability as to its conclusions so that firms have sufficient recourse for instances of claimed false positives.

Financial institutions seeking to employ AI models in AML and KYC compliance processes should also carefully monitor all information fed into the model, whether for model training or for model decision-making purposes. To the extent any outputs from an AI model might be made available outside of the firm, information fed into the model should exclude any confidential information of the firm and of its customers to ensure compliance with both confidentiality obligations and applicable data privacy requirements imposed on financial institutions.

II) AI KM – Financial Services Consumer Banking Fraud Detection and Prevention

Wherever financial services providers enable consumer transactions, the risk of fraud is a central issue that must be addressed and mitigated. The rise of technology-enabled digital payments has created an arms race between increasingly prevalent fraudsters, and financial services providers utilizing sophisticated tools to detect and deter fraud. Artificial intelligence has accelerated this trend: **predictive AI** has empowered financial services providers to better detect fraud in real time, allowing them to decline potentially fraudulent transactions before they are processed, while **generative AI** is now enabling fraudsters to more efficiently and more effectively fight through these defenses. In this section we will explore the use of AI in consumer transaction fraud detection and protection, outlining the current technological landscape and how this may evolve going forward.

II.I) HISTORY OF FRAUD DETECTION TECHNOLOGY

The increasing prevalence of digital payments in today's economy has invited an uptick in consumer banking fraud. According to the Federal Trade Commission, in 2022 consumers reported losing approximately **$8.8 billion** to fraud—up more than **30%** from 2021. (link). Fraudsters who once needed physical access to credit or debit cards to perpetrate fraud can now target a wide variety of security vulnerabilities across a breadth of digital payment and e-commerce platforms. Historically, financial services providers have taken rigid, rules-based approaches to detecting and preventing payments fraud. For instance, attempted payments would be flagged as potentially fraudulent based on geographic location, payment amount, payment time, or other pre-determined limits. Such legacy systems had critical inherent shortcomings: they failed to adapt to changing spending habits of consumers over time without costly and time-consuming manual updates, and they could be learned (and avoided) by practiced fraudsters. These shortcomings led to high rates of false negatives, as a high volume of fraudulent transactions slipped through the cracks.

The early adoption of AI-enabled fraud detection tools helped financial service providers level-up in their efforts to deter fraud. In many ways, the questions involved and the data sets available to such financial services providers are ideally suited to the application of predictive AI, making use of pattern recognition across large data sets. Whether any given transaction is fraudulent can be predicted with relative confidence based on how well it fits into past patterns of payments known to be legitimate. Banks already had large volumes of prior transactions data—including pre-labelled fraudulent transactions data derived from legacy manual and rules-based fraud detection efforts—on which they could train AI models. Furthermore, financial services providers could effectively lean on their customers to provide additional reinforcement learning for predictive AI models in the form of real-time email and text message-based suspect transaction validation. Thus, once the underlying technology sufficiently matured, it was comparatively (relative to other applications in other industries) easy for financial services providers to deploy in the name of fraud detection.

II.II) ISSUES ARISING IN ADOPTION

That is not to say early adoption of predictive AI in the fraud detection space was without issue. Heightened monitoring of large data sets by generalized predictive AI models has shifted the most common source of fraud detection error from false negatives to false positives). This problem was likely compounded by the low risk tolerances of many financial services providers, who, when determining their desired sensitivity of fraud detection AI models preferred to err on the conservative side of enhanced caution. Said differently, financial services providers may have realized the asymmetric cost-benefit balance of false positives, and calibrated their fraud detection models accordingly: in the era of near instantaneous text messages compounded with the low risk tolerance of many financial services providers transaction validation, the cost to consumers of false positives (measured in seconds of inconvenience) is small compared to the cost to financial services providers (measured in dollars lost processing fraudulent transactions) of false negatives. Additionally, more heavily data-driven approaches to fraud detection have raised questions related to cybersecurity and privacy, as some critics question how much data is worth sharing in the name of deterring fraud. Questions of potential bias, transparency, and fairness linked to the black box nature of underlying AI models likewise abound. Clearly there remains much room for improvement in the way financial services providers design, train, deploy, and calibrate AI based fraud detection tools in the consumer financial services space.

II.III) LOOKING TO THE FUTURE

Looking forward, we foresee two main avenues for improvement of the use of AI in consumer financial services fraud detection. The first represents a change in degree from existing applications, and is likely to occur in the near term. As financial services providers gain access to more consumer data and more processing power, they will be better able to tailor fraud detecting AI models to narrower subsets of consumers—or perhaps even individual consumers—yielding more accurate analyses of each transaction in the context of the corpus of those consumer's transactions history. Incorporating an individual's cell phone location data into predictive AI models, for example, could drastically improve such models' ability to predict the validity of a transaction. Today, some companies (e.g., Sardine) are even going far beyond this and checking against how a phone is being held, and other data that can greatly impact fraud detection.

However, financial institutions will need to keep in mind the privacy implications of using such data and whether additional disclosures need to be made or whether new consents will be required. In addition, firms will need to consider whether use of new information leads to biased or unreliable results, by for example penalizing customers whose travel schedules suddenly change.

The second expected change represents a change in kind from existing AI applications, and is likely to take longer to develop. Financial services providers may look beyond supervised learning AI models based on pre-labeled data and towards unsupervised models, and use of unstructured data, to expand the scope of fraud detection capabilities and reduce the need for human input. Generative AI could enable financial services providers to analyze unstructured data and interact more meaningfully with clients—improving efficacy along the way.

III) AI Standards – Credit Decisions

Traditionally, lenders had only limited data to determine the creditworthiness of an individual, such as debt, income, and loan payment history. AI and "big data" have exponentially expanded the factors available to inform credit decisions, allowing lenders to issue loans to "credit invisibles," or those without extensive debt, income, or loan payment history. However, this new data also risks exposing lenders (and companies providing such data) to increased regulatory oversight. This section highlights the legal considerations when using "big data" in credit decisions, and provides a few suggestions to ensure companies do not inadvertently expose themselves to greater legal risk.

III.I) WHAT TYPES OF DATA ARE USED IN CREDIT DECISIONS?

In the realm of credit decision-making, there is an ongoing and notable shift from traditional methods to a technologically-driven, data-rich approach. This transformation is facilitated by the **integration of AI** and the increasing availability of alternative data sources. The traditional or "classic" data that is utilized when determining an individual's creditworthiness includes factors such as their FICO score, debt levels, income, and credit history (including credit card usage, auto and personal loans, and mortgages, among others). These inputs have long been the cornerstone of credit assessments, providing a snapshot of an individual's financial stability and reliability.

However, the advent of AI and the vast amount of data generated in our increasingly digital world have opened the doors to an array of new data sources for credit assessment. This expanded dataset includes education information, address stability, rent and utility payment history, online shopping activity, browsing history, and even inferences drawn from this data, such as detecting signs of marriage infidelity.

Social media activity, phone apps downloaded, standardized test scores (like SAT), GPA, field of study, job history, geolocation data, payday loan usage, bank account balances, student loan debt, and even smartphone usage patterns, such as the time of day calls are made, the length of phone calls, texting frequency, text length, phone make and model, phone contact organization, Wi-Fi networks used, mobile wallet balances, and phone battery level trends are all now on the radar of AI-assisted credit assessment. In some cases, even one's friends and contacts, along with their credit and personal information, can be considered. Type of computer used and email domain are additional data points that can be of influence.

While this extensive array of data offers the potential for more accurate credit assessments, it raises substantial **privacy concerns**. For example, recent polls indicate that **96%** of respondents are opposed to the use of social media data for credit risk assessment.[10] The broad spectrum of data inputs noted above is certain to raise eyebrows from a privacy standpoint, as it essentially opens up individuals' personal lives to be evaluated in creditworthiness assessments. The ethical implications of this vast data collection and its use in determining an individual's creditworthiness are profound, and they underscore the necessity of robust data protection and privacy regulations.

There are also significant legal considerations under the **Fair Credit Reporting Act** (FCRA). The FCRA governs credit reporting agencies, which play a vital role in credit decisions. As data aggregators provide increasingly detailed "profiles" to employers, creditors, and similar entities for the purpose of informing credit decisions, there is a growing risk that these aggregators could be categorized as "credit reporting agencies."[11] In such a scenario, they would be subject to the full scope of the FCRA, including its stringent requirements for data accuracy, consumer rights, and dispute resolution. This legal dimension raises questions about the regulatory framework and potential consequences for the industry.

III.II) ECOA: DISCLOSURE & DISCRIMINATION

Under the **Equal Credit Opportunity Act** (15 U.S.C. §1691) ("ECOA"), creditors are mandated to furnish a written statement to applicants outlining the specific reasons for taking adverse actions, such as refusing a loan application. These reasons must not only be relevant but also accurately represent the factors that the creditor genuinely considered or assessed in their decision-making process. It is crucial to note that no factor deemed a primary basis for the adverse action can be omitted from the disclosure. The Consumer Financial Protection Bureau (CFPB) has emphasized that "creditors cannot justify noncompliance with ECOA based on the mere fact that the technology they use to evaluate credit applications is too complicated, too opaque in its decision-making, or too new."[12] This underscores the importance of transparency and accountability in the use of technology for assessing credit applications, ensuring fair treatment for all applicants under the ECOA.

III.III) ECOA: DISCRIMINATION

The **Equal Credit Opportunity Act** (ECOA) serves as a fundamental safeguard against discrimination in credit transactions. The ECOA explicitly forbids creditors from engaging in discriminatory practices against credit applicants based on several criteria, including race, color, religion, national origin, sex, marital status, age, the receipt of income from a public assistance program, or the exercise of any right under the Consumer Credit Protection Act.

However, with the advent of AI-driven credit models, some of the data points considered in these models can sometimes act as proxies for characteristics such as race, religion, or sex. Consequently, the weight and scoring applied by AI-driven credit models may inadvertently result in proxy discrimination, where "the predictive power of a facially-neutral characteristic is at least partially attributable to its correlation with a suspect classifier," potentially undermining the core principles of ECOA.[13] This issue raises important questions about fairness and bias within the framework of AI-assisted credit assessment.

III.IV) DODD-FRANK: UNFAIR ACTS OR PRACTICES[14]

Under the **Dodd-Frank Wall Street Reform** and **Consumer Protection Act of 2010** (124 Stat. 1376) ("Dodd-Frank"), it is unlawful for any provider of consumer financial products or services to engage in any unfair, deceptive or abusive act or practice ("UDAAPs"). The CFPB is tasked with protecting consumers against such UDAAPs. Any entity that determines creditworthiness when issuing loans would quality as a provider of consumer financial products, subject to CFPB jurisdiction. Under Dodd-Frank, an act is "unfair" when (i) it causes or is likely to cause substantial injury to consumers; (ii) the injury is not reasonably avoidable by consumers; and (iii) the injury is not outweighed by countervailing benefits to consumers or to competition.

A company that uses AI-driven algorithms to make a credit decisions may inadvertently commit an "unfair" act under Dodd-Frank because (i) a denial to credit could cause substantial injury to consumers, (ii) without access to the model, the injury would not be avoidable by the consumer, and (iii) depending on the accuracy of the model, might not be outweighed by countervailing benefits to consumers.

The crux of the issue turns on prong (iii). On the one hand, AI-driven credit decisions may increase access to credit to "credit invisibles", or those without a traditional credit score (i.e., one driven by debt, income, and assets, as noted above). On the other hand, such tools may also exacerbate discriminatory results in credit decisions (e.g., through proxy discrimination, as noted above).

III.V) DODD-FRANK: AVM RULES[15]

Federal agencies recently proposed rules that require banks, when using automated valuation models ("AVMs") in mortgage decisions, to adopt policies/procedures designed to, among other things, (i) ensure a high level of confidence in the estimates produced by AVMs; (ii) promote compliance with applicable nondiscrimination laws; (iii) avoid conflicts of interest, and (iv) protect against the manipulation of data. If credit issuers cannot understand or articulate their model outputs, then banks risk noncompliance with these (proposed) rules.

AUTONOMOUS VEHICLES USE CASES

Autonomous vehicles (AVs) represent a transformative approach to transportation, leveraging advanced sensors, artificial intelligence, and connectivity to operate without human intervention. These vehicles are expected to bring about new transportation use cases influenced by factors such as the type of cargo, ownership models, and operational environments. While the potential of AVs is vast, achieving true autonomy, where no human intervention is required under any circumstances, remains a challenge. The integration of technologies like 5G, edge computing, and vehicle-to-everything (V2X) communication will be pivotal in realizing the full potential of AVs in the future.

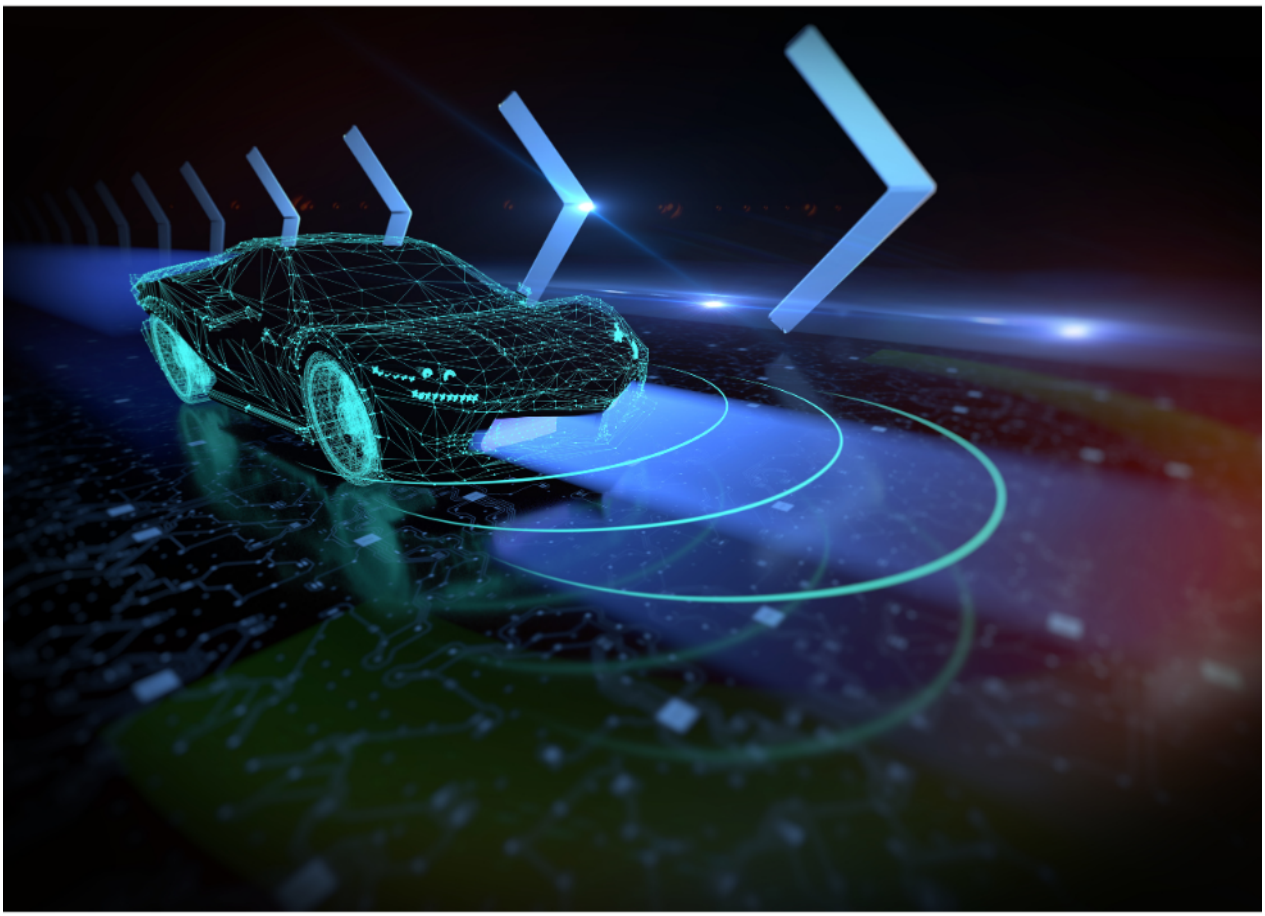
Autonomous vehicles can be much more widely interpreted to include drones (air and sea), as well as any other vehicles equipped with computer vision and AI. Equipping any vehicle with vision and AI-based interpretation will change not only land and air transport but also manufacturing with robots and armed conflicts. A fundamental aspect of AI is to build machines that can simulate humans in their operations in physical spaces – such as factories or office functionalities taking place out of new physical spaces.

AI, particularly predictive AI, has the potential to enhance the benefits of computer vision based on massive data and tailor-made hardware from cameras to sensors and AI chips. Computer vision and AI will let machines and vehicles operate and interact with our real-world physical environment. Many US AI companies are investing significantly in this sector.

Companies and organizations venturing into the realm of autonomous vehicles (AVs) are making significant strides in both technology development and deployment. Here are some notable examples: Microsoft, Alphabet, Baidu, General Motors Company, NVIDIA, Tesla, Ford, Aptiv PLC, Luminar Technologies, Pony.ai, and others.

Companies and organizations looking to implement autonomous vehicle (AV) technologies should consider several critical factors. **The potential of autonomous driving (AD) to transform transportation, consumer behavior, and society is vast**. However, to realize the consumer and commercial benefits of AD, auto OEMs and suppliers may need to develop new sales and business strategies, acquire new technological capabilities, and address concerns about safety. Challenges such as object detection, decision-making, and handling edge cases are paramount. Furthermore, testing and validation in the AV realm will require a paradigm shift. Instead of relying solely on physical testing, companies will need to adopt software-based simulations and virtual testing methods to ensure the safety and reliability of AV systems. As the technology evolves, addressing these concerns will be crucial for the mass adoption and success of AVs in the market[16].

In addition, companies and organizations delving into the realm of autonomous vehicles (AVs) must be attuned to the regulatory landscape, and necessary safeguards required across various jurisdictions. The adoption of AVs hinges on global regulations that favor both testing and development, ensuring the safety of all road users. Regulatory bodies, such as the United Nations Economic Commission for Europe (UNECE), alongside several countries, are striving to refine a global regulatory framework that addresses the multifaceted challenges posed by AVs. The overarching goal is to strike a balance between innovation and safety, ensuring that as AVs become more prevalent, they do so in a manner that instills trust and confidence in the public.



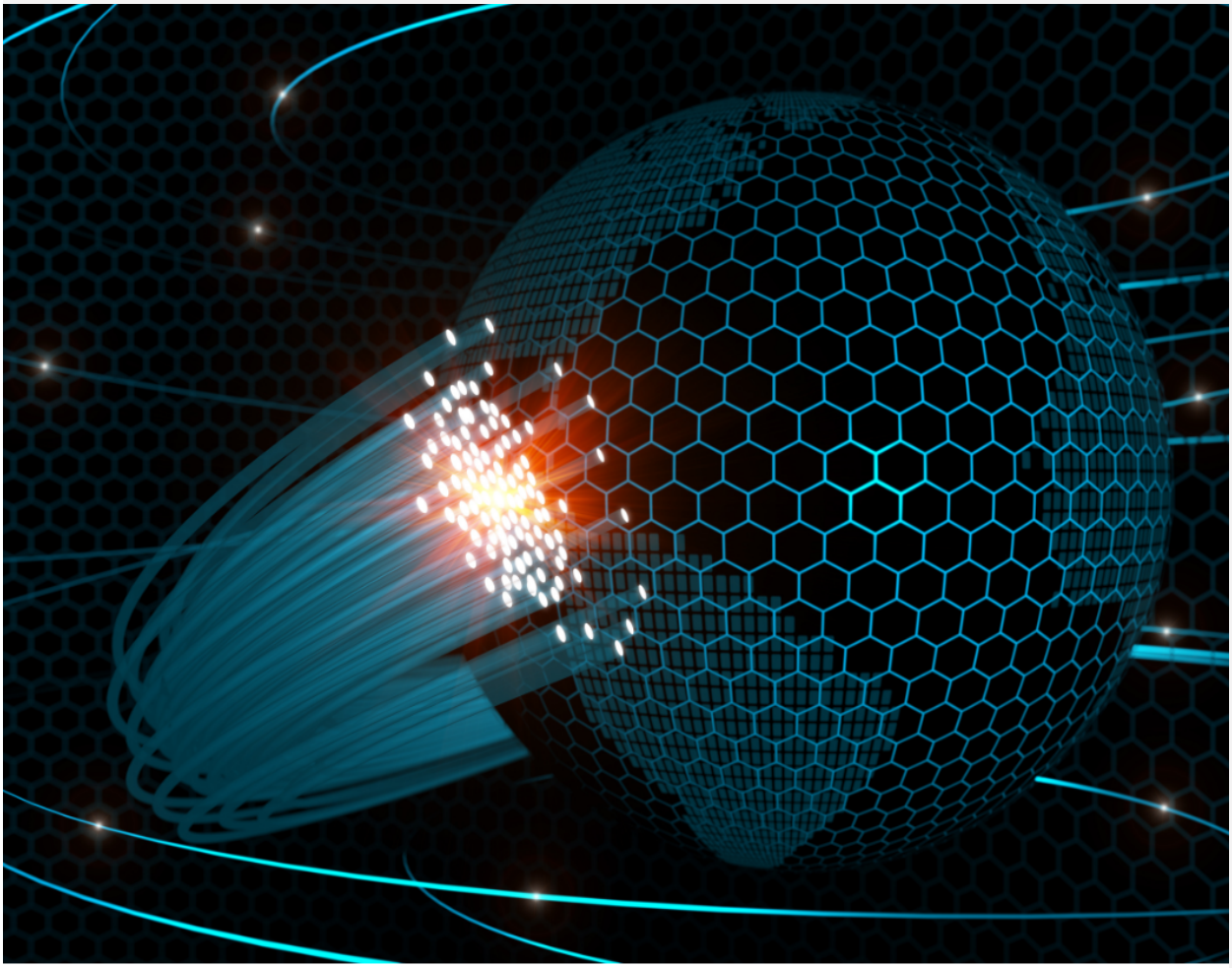
McKinsey's report *"Global Autonomous Vehicles Regulatory Growth Opportunities"* offers insights into testing and deployment regulations for autonomous vehicles in **27** countries, covering regions like Europe, Asia-Pacific, and North America, including China. In addition, the US and EU have held consultations on the matter.

Among these countries, Germany, China, and Japan are among the early actors at the forefront of crafting a regulatory framework conducive to the evolution and deployment of AVs. Germany, a hub for several automotive powerhouses, has a national strategy[17] in place, which intends to set international standards automated and connected driving systems to perform their functions safely and reliably, across national boundaries, while ensuring clear regulation for rights to individual mobility data. The country is diligently working to expand these frameworks for broader applicability. China, another early player in AV testing, has not only implemented comprehensive road safety laws covering driverless vehicles but also facilitated local governments to introduce their bespoke regulations. At the national level, Chinese authorities have rolled out Regulations on the Administration of Road Testing of Autonomous Vehicles[18], a pivotal step to foster transportation innovation and ensure the safe integration of AVs on roads. Japan also adopted provisions for automated driving[19] in April of 2023, to ensure safe and early deployment of automated driving systems in accordance with the existing safety frameworks.

GOVERNMENT & POLICY USE CASES

The government's place is not in the development of AI, but in providing governance to allow public servants to use it to better serve their constituencies. The integration of AI in various sectors, especially government, presents multifaceted challenges and opportunities. Traditional forms of service provision, policy-making, and enforcement are undergoing rapid transformations with the introduction of AI technologies. Governments worldwide are recognizing the potential of AI to revolutionize public-sector ecosystems, but this also brings forth complexities in terms of implementation, transparency, and accountability. The expanding use of AI in governance can significantly alter the dynamics of public service delivery and decision-making processes.[20] In the United States, the state of Ohio has explored using a large language model (LLM), an AI program that can perform tasks such as recognizing and generating text.

AI, especially when deployed in public governance, can inadvertently introduce biases and discriminatory decisions. Policymakers are increasingly focused on the risks associated with AI technologies making discriminatory decisions, similar to human biases. These biases can stem from the data sets on which AI models are trained or from the algorithms themselves. There have been instances, particularly with facial recognition software, where misidentification of individuals, especially those in minority groups, has raised concerns. To address these challenges, there's a pressing need for policies that ensure AI systems are transparent, fair, and free from biases. A report from the National Institute of Standards and Technology (NIST) emphasizes the importance of mitigating biases through appropriate representation in AI data sets and rigorous testing and validation of AI systems.[21]



Immediate assessment regarding AI implementation must include current use of AI as well as automation intelligence by the organization, to understand what is already being utilized and the processes that govern them. Once an audit is completed, a governance structure should be put in place with leadership from across the government and organization, with particular attention paid to representatives from the procurement space that will have to acknowledge and plan for a new element to procuring artificial technology. The structure should include core values by which every new artificial intelligence tool must abide, a mandate for the governing body to create a regulatory process for internal use, and a direction to design updated procurement processes allowing for the accurate procurement of AI technology.

Drawing parallels with AI, other technological advancements in the past have also posed challenges that required regulatory and policy interventions. For instance, the internet's advent brought about issues related to data privacy, cybersecurity, and digital rights. Over time, governments and organizations established frameworks and guidelines to address these concerns. Similarly, AI's integration in public governance can benefit from lessons learned from these previous technological shifts. Policymakers can look at successful regulatory models from other tech domains and adapt them to the unique challenges posed by AI. By doing so, they can ensure that AI is harnessed responsibly and ethically, maximizing its benefits while minimizing potential harms.

POLICY IMPLICATIONS OF USE CASES

From a policy perspective, it is important to consider the ways AI raises complex issues, including ethical issues, and ways to address them. Policymakers are beginning to evaluate measures than need to be taken to address the risks and novel issues that AI poses. For instance, the US White House released a pioneering and very comprehensive policy statement with the "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence"[22] that attempts to respond to the entirety of treats and corresponding challenges posed by AI. The states of Pennsylvania and Virginia have also produced executive orders on AI. AI policies have also been envisioned in Japan, China, and India.

With respect to security and privacy, applicability of data protection regulations and safe data handling practices are key. To safeguard against security threats, particularly at a geopolitical level, governments and organizations internationally must agree on best practices for AI developments and deployments, with underlying international cooperation toward global norms and regulations.

Policies and other initiatives should also promote economic equality with AI developments – including reskilling programs for sow skilled workers, promoting social safety nets, and fostering inclusive AI developments that can enable more equal opportunities that can combat economic inequalities rather than aggravate them.

Moreover, many of the proposed measures by policymakers to address AI risks can also be relevant for blockchain developments such that, in convergence with AI, can support innovation to benefit human civilization.

Importance of Principles & Standards

In the immediate term, principles, globally agreed upon regarding AI, offer a softer approach that sets the stage for more hard core policies to be enforced in the future. For example the OECD AI Principles were adopted in 2019 by member countries, followed then by an adoption by the G20, giving them a global reach with the 2 AI super powers (US and China) agreeing on them. Principles contain Ethical considerations to ensure that the use of these technologies aligns with societal values and does not result in discrimination, bias, or harm to individuals or groups.

In addition, voluntary standards, can be made obligatory through later regulations. For AI, the standards bodies the International Organization for Standardization (ISO) and the Institute of Electrical and Electronics Engineers Standards Association (IEEE) are actively working on developing standards around development and deployment of AI tools, such as IEEE's seminal work on Ethically Aligned Design.

FAT/ML, or Fairness, Accountability, and Transparency in Machine Learning, is an interdisciplinary field of research and practice that focuses on addressing ethical and social concerns related to machine learning and artificial intelligence (AI) systems. FAT/ML encompasses several key principles and areas of focus:

- **Fairness -** This aspect of FAT/ML aims to ensure that machine learning algorithms and AI systems do not discriminate against or unfairly disadvantage certain groups of people. Fairness concerns often involve issues related to bias in data, algorithmic decision-making, and the potential for reinforcing or exacerbating existing societal inequalities.
- **Accountability -** Accountability in FAT/ML refers to the ability to trace and attribute decisions made by AI systems to specific individuals or entities. This involves understanding how decisions were reached, what data was used, and who is responsible for the outcomes. Accountability mechanisms help establish transparency and ethical responsibility.
- **Transparency -** Transparency involves making AI and machine learning models more understandable and interpretable. This is important for both technical experts and non-experts to comprehend how algorithms work, what factors influence their decisions, and how to assess their behavior.
- **Privacy -** FAT/ML also considers the protection of individual privacy in the context of AI and machine learning. It involves implementing measures to safeguard sensitive and personal information, ensuring that data is used responsibly and in compliance with data protection laws and regulations.
- **Robustness -** Ensuring that AI systems are robust to adversarial attacks and unexpected inputs is another aspect of FAT/ML. Robustness measures aim to prevent AI systems from making incorrect or harmful decisions when faced with unusual or malicious inputs.

FAT/ML often intersects with discussions about regulation and policy development for AI and machine learning. Researchers and policymakers work together to establish guidelines and rules that promote fairness, accountability, transparency, and ethical behavior in AI applications.

Government Actions and Public LLMs

Regulatory Clarity

Governments must help provide clear and up-to-date regulations and guidelines for the responsible use of AI and blockchain. This clarity will in turn help organizations make informed decisions and comply with the law. The issue is that AI is a very fast moving field while laws and regulations are slow, adding to that the lack of comprehension by policy makers of the impact of these systems and how they operate. A robust regulatory framework is paramount. This includes clear guidelines on testing protocols, safety standards, data privacy, and interoperability.

Open Source vs. Closed Source

The choice between open-source and closed-source LLMs depends on the specific use case and requirements. Open-source models promote transparency and collaboration but may require more effort to customize and maintain. Powerful Open-source models could also be used for nefarious reasons and that is why many are suggesting a graduation approach to decide what model should be open or close and for what purpose, which user, etc. Closed-source models offer proprietary features and support but are by definition less transparent.

Public LLMs

Governments may consider encouraging or supporting the development and use of public LLMs for various purposes, such as legal research, content generation, and more.

Monitoring and Oversight, Collaboration

Governments are looking at establishing mechanisms for monitoring AI and blockchain implementations, especially in critical sectors like healthcare, finance, and transportation, to ensure compliance with regulations and ethical standards. Governments should also facilitate collaboration between industry stakeholders, academia, and civil society to develop best practices, standards, and frameworks for AI and blockchain governance.

CONCLUSION

AI raises complex issues, especially as a technology enabler in social contexts with increasing levels of nuance. Therefore it is critical to consider the impact of AI on human society and well-being. In order to comprehend the impact of AI and promote its adoption in ways that are beneficial for human civilization, cooperation among stakeholders is key – an effort that will likely be driven by standards and regulation. Now more than ever, cooperation among stakeholders is essential to advance harmonized regulations for coordinated and constructive AI innovations that will benefit humanity. Agreement on standards, conditions, and parameters will shape the future of AI and its impacts.

There is a need to balance technological development while preserving the integrity of human interactions, in order to maintain our well-being and flourishing. AI developers and researchers must engage in robust testing, validation, and monitoring of AI systems to identify and address any issues and unintended consequences before they escalate. The AI community must also promote safety research, ethical guidelines, and transparency, in particular for artificial general intelligence developments, such that they can serve humanity's best interests rather than posing serious threats.

In this context, we believe that blockchain technology has an important role to play in the development of responsible AI. Blockchain can secure, source, and verify data provenance, for a future landscape of AI that is made more trustworthy. Blockchain serves as a risk mitigation tool, with a transparent ledger and audit system that support an unprecedented level of effective and trustworthy record keeping. Blockchain-based identification mechanisms can address many of the privacy concerns that arise from Machine Learning. Digital rights will be a foundational piece of this.

We hope to spur the first major body of work to explore the convergence of blockchain and AI, with continued collaborations and discussions toward responsible innovation.

ANNEX A

MORPHOLOGICAL FRAMEWORK OF AI

The morphological framework below, based on Fritz Zwicky's "Morphological Astronomy," is a proposal to discuss and agree upon dimensions and conditions for AI implementations. This is a mind map of detailed categorizations and sub-categorizations of AI features based on structures, types, tasks, functions, and branches, to provide a broad structure that can be useful multiple stakeholders and facilitate collaboration.

ENDNOTES

AI & CONVERGENCE

1 https://creativecommons.org/2023/08/18/understanding-cc-licenses-and-generative-ai/
2 https://arxiv.org/pdf/2305.03928.pdf
3 Refer to taxonomy resources of GSMI 4.0, which includes a subset of definitions on AI
4 https://evals.alignment.org/
5 https://www.data4sdgs.org/
6 12 CFR Part 326 (https://www.ecfr.gov/current/title-12/chapter-III/subchapter-B/part-326); FinCEN (https://www.fincen.gov/resources/statutes-and-regulations/bank-secrecy-act#:~:text=Specifically%2C%20the%20regulations%20implementing%20the,might%20signify%20money%20laundering%2C%20tax)
7 12 CFR Part 353 (https://www.ecfr.gov/current/title-12/chapter-III/subchapter-B/part-353)
8 CDD Rules (https://www.federalregister.gov/documents/2016/05/11/2016-10567/customer-due-diligence-requirements-for-financial-institutions)
9 FinCEN (https://www.fincen.gov/resources/statutes-and-regulations/cdd-final-rule)
10 https://www.consumer-action.org/news/articles/alternative_data_and_financial_inclusion_summer_2017
11 https://www.ftc.gov/business-guidance/blog/2012/06/speaking-spokeo-part-1
12 https://www.consumerfinance.gov/about-us/newsroom/cfpb-acts-to-protect-the-public-from-black-box-credit-models-using-complex-algorithms/
13 https://www.brookings.edu/articles/credit-denial-in-the-age-of-ai/
14 https://www.cfpaguide.com/portalresource/Exam%20Manual%20v%202%20-%20UDAAP.pdf
15 https://occ.gov/news-issuances/bulletins/2023/bulletin-2023-21.html
16 https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/autonomous-driving-disruption-technology-use-cases-and-opportunities
17 https://bmdv.bund.de/SharedDocs/EN/publications/strategy-for-automated-and-connected-driving.pdf?__blob=publicationFile
18 https://www.mps.gov.cn/n2254536/n4904355/c7787881/content.html
19 https://www.npa.go.jp/english/bureau/traffic/selfdriving.html
20 https://www.sciencedirect.com/science/article/pii/S0740624X21000137
21 https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/autonomous-driving-disruption-technology-use-cases-and-opportunities
22 https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/#:~:text=It%20is%20necessary%20to%20hold,equity%2C%20and%20justice%20for%20all

GLOBAL BLOCKCHAIN BUSINESS COUNCIL

DC Location:
1629 K St. NW, Suite 300
Washington, DC 20006

Geneva Location:
Rue de Lyon 42B
1203 Geneva
Switzerland