IDVerse™

Ver. 1.0 2023

PUBLIC SECTOR

The Regulatory Path Forward\

Driving AI Oversight to Keep Up with AI Advancement

AN ETHICAL FRAMEWORK
For Eliminating Bias and Promoting Inclusion and Fairness

THE ROLE OF GOVERNMENT
In Managing Generative AI and Remote Identification Verification Technology

ZERO BIAS AI
Frameworks and checklists to address algorithmic bias

Table of Contents

Introduction

In an economy where online transactions are beginning to supersede in-person interactions, identity verification (IDV) technology has become the critical connective tissue between businesses and consumers everywhere. It's now considered table stakes in almost every industry, from banking to employment to gaming.

The most advanced IDV solutions use generative AI and neural networks along with biometric data, such as a person's face, to authenticate users seeking to perform a high-risk activity like withdrawing funds from a bank account or placing a sports bet. One requirement, therefore, is to create solutions that stay ahead of bad actors by using this technology better than they do.

Synthetic media, more commonly known as "deepfakes," enable the creation of highly realistic, manipulated video, image, and audio hoaxes that can convincingly impersonate individuals. They can be incredibly hard to detect with the unaided human eye. In the commission of deepfake-facilitated identity fraud, perpetrators can gain access to resources, commit crimes, and damage reputations.

But institutions face a larger threat than fraud and deepfakes: the pervasive issue of AI bias. Ensuring that biometric and document verification technologies function effectively for users regardless of ethnicity, age, sex, or gender identity is vital.

In two parts, this report explores the need for guiding principles and regulation—both embedded within the processes of AI enterprises themselves and externally from government entities—to ensure that decision-making systems are transparent, accountable, and subject to oversight:

1. An Ethical Framework for Eliminating Bias
2. The Role of Government in Managing AI

Part 1: An Ethical Framework for Eliminating Bias

Bias can manifest in a myriad of ways. For instance, a seemingly innocuous prompt such as "show me faces of successful business owners" on a typical image-generative AI platform like Midjourney can disproportionately favor Western, white, male faces, neglecting the fact that the vast majority of the world's population—more than 75%—lives in Asia and Africa.

Sydney Harbor Bridge
The bridge has around 6 million rivets, which means there are more rivets on the bridge than there are people in Sydney.

AI programs often reinforce stereotypes because human engineering teams have inherent biases that are encoded into the algorithms. As leaders of technology, we have therefore found ourselves at the crossroads of innovation and responsibility. Our creations—the software solutions we engineer—can either exacerbate societal biases or serve as a force for positive change; the choice is ours to make.

Pledging an oath

To face this challenge head on, IDVerse is introducing the Code Zero Bias Oath, inspired by the enduring principles of the Hippocratic Oath, but tailored to the unique challenges and opportunities of our field. This oath embodies our collective commitment to reducing algorithmic bias, promoting fairness, and upholding the highest ethical standards in AI software development:

Code Zero Bias Oath

I, as a software engineer, solemnly swear to uphold the principles and practices outlined in this Code Zero Bias Oath. In my pursuit of designing, developing, and deploying software, I commit to the following:

1. ***Do No Harm**: I shall prioritize the well-being of individuals and communities who may be affected by the software I create. I will strive to ensure that my work does not cause harm or perpetuate bias, discrimination, or inequality.*
2. ***Equity and Fairness**: I will actively seek to identify and rectify biases in algorithms and data sets. I pledge to promote fairness and impartiality, striving to create software that treats all individuals equally regardless of their background, race, gender, or any other characteristic.*
3. ***Transparency and Accountability**: I will be transparent about the decision-making processes and data sources used in my software. I accept responsibility for the consequences of my work and will be accountable for any biases or ethical lapses that may arise.*
4. ***Inclusivity**: I will advocate for diverse and inclusive teams, recognizing that different perspectives lead to more robust and ethical solutions. I will actively work to create an environment where underrepresented voices are heard and valued.*

5. ***Continuous Learning***: *I understand that technology evolves rapidly, and I commit to staying informed about emerging best practices, guidelines, and regulations related to algorithmic bias and ethical software development.*
6. ***User Privacy and Consent***: *I will respect user privacy and seek informed consent for data collection and usage. I will implement strong data protection measures to safeguard user information.*
7. ***Mitigation and Remediation***: *If I discover bias or ethical concerns in software I have developed, I will take immediate steps to mitigate harm and rectify the issues. I will report such concerns to relevant stakeholders and take corrective action.*
8. ***Community Engagement***: *I will actively engage with the communities impacted by my software, seeking their feedback and addressing their concerns. I will be open to criticism and commit to improving my work based on community input.*
9. ***Regulatory Compliance***: *I will adhere to all relevant laws, regulations, and industry standards related to algorithmic fairness and data ethics in software development.*
10. ***Advocacy for Ethical Technology***: *I will advocate for the responsible and ethical use of technology within my organization and the broader industry. I will use my influence to promote ethical practices and raise awareness about the importance of reducing algorithmic bias.*

I acknowledge that my work as a software engineer has a profound impact on society, and I accept this oath as a solemn commitment to ethical software development. I will strive to uphold these principles throughout my career, recognizing that my actions can shape the future of technology and its impact on humanity.

By taking this Code Zero Bias Oath, software engineers demonstrate their dedication to ethical software development,

with a focus on reducing algorithmic bias and promoting fairness, transparency, and accountability.

Nelson Mandela Bridge
This Johannesburg landmark symbolically links the old and new as it ushers traffic into the heart of rejuvenated downtown Johannesburg, and into Newtown from Braamfontein.

Training data & fair-sourced faces

A flawed algorithm is not the only source of bias that can undermine the fairness and accuracy of an identity verification system. If biased data is used to train the algorithm, it is likely the system will exhibit those same biases when making decisions.

In the context of facial IDV systems, “training data” refers to specialized datasets of facial images used to train the machine learning algorithms or deep neural networks that are responsible for recognizing and verifying individuals' identities based on their facial features.

Here's how training data works in facial identity verification:

- **Features**: The training data includes a vast collection of facial images as its primary feature. These images represent various individuals, captured under different lighting conditions, angles, and backgrounds. Each image provides information about the facial characteristics of a person, including the position of facial landmarks (such as eyes, nose, and mouth), skin texture, and other unique details.

- **Labels or ground truth**: In supervised learning for facial identity verification, each facial image in the training data set is associated with a label or ground truth, which specifies the identity of the person in the image. These labels serve as the correct reference for the algorithm during training. For instance, a label might indicate that a particular image is of "John Doe."
- **Training process**: The machine learning or deep learning algorithm uses this labeled training data to learn and extract relevant patterns and features from facial images. It analyzes the unique characteristics of each individual's face, such as the arrangement of facial landmarks, and learns how to distinguish one person from another.
- **Model development**: Through the training process, the algorithm adjusts its internal parameters to minimize the difference between its predictions and the ground truth labels in the training data. This fine-tuning allows the model to become increasingly accurate in recognizing and verifying identities based on facial features.
- **Generalization**: Once the algorithm has been trained on the dataset, it can be deployed for real-time facial identity verification tasks. The algorithm applies the knowledge it acquired from the training data to analyze and compare facial features in new, unseen images to verify the identity of individuals.

The quality and diversity of the training data determine the performance and efficacy of a facial identity verification system. In particular, training data affect the fairness of decision-making algorithms and their ability to mitigate bias, or achieve Zero-Bias AI™.

Training data should encompass a wide range of facial variations—including age, gender, ethnicity, and lighting conditions—to ensure that the algorithm can accurately identify individuals from different backgrounds and under different circumstances or environments.

Ethical procurement & data nutrition labels

With face recognition technology becoming increasingly prevalent, the ethical sourcing of face biometrics data used to train identity verification systems should be a top concern for businesses and their customers. Training data sets should have the equivalent of a “nutrition label” which would provide crucial information about the dataset's sources and characteristics, helping developers make informed decisions.

Ethical procurement in face biometrics technology centers around ensuring that the data used to train and develop face recognition systems adheres to strict fairness and ethical guidelines. This not only safeguards individual privacy and civil liberties, but also ensures that the technology is as accurate and fair as possible.

Key considerations for ethical procurement

Here are some principles to keep in mind when considering the data an IDV provider has used for the training of its machine learning system:

- **Transparency**: Transparent providers are open about their data collection methods, sources, and usage. They are willing to disclose their practices and allow third-party audits to ensure ethical compliance.
- **Consent**: Ethical providers obtain explicit consent from individuals whose data is used for training purposes. They make it clear how the data will be used and for what purposes. Further, consent must be permissible under local regulations. As an example, GDPR regulations preclude a remote ID verification vendor (as data processor) from receiving consent from an end-user to use his/her data for training; that consent remains in the strict purview of only the data controller.
- **Data minimization**: Ethical procurement minimizes data collection to only that which is necessary for the

technology's intended purpose. It avoids the unnecessary collection and storage of sensitive data.

- **Fair representation**: Providers ensure that their training datasets are diverse and representative of the population, minimizing biases and ensuring equitable performance across demographic groups.
- **Security**: Ethical providers prioritize robust security measures to protect the data they collect. They follow best practices for data encryption, access control, and regular security audits.
- **Data deletion**: Providers have clear policies and procedures for data retention and deletion. They respect individuals' right to have their data removed upon request and have system capabilities for business buyers to execute desired privacy policies.

Millennium Bridge
The London Millennium footbridge is the first pedestrian bridge to cross the River Thames in over 100 years.

Looking ahead

As facial recognition technology continues to evolve, ethical considerations must remain at the forefront of businesses and consumers alike. Ethical procurement of data ensures that this technology is used responsibly, respects individual rights, and remains an asset to society rather than a threat to privacy and civil liberties.

Organizations must set a positive example and contribute to the development of a more responsible and ethical landscape for face biometrics technology. Businesses, suppliers, and consumers alike should prioritize transparency, consent, data minimization, fairness, security, and adherence ethical standards.

Ethical procurement of face biometrics training data is not just a corporate responsibility; it's a societal imperative in an age where facial recognition touches so many aspects of our lives.

Part 2: The Role of Government in Managing AI

We have reached a societal tipping point when it comes to artificial intelligence. Technologies that were once only theoretical are now reality—self-driving cars navigating our city streets; computer vision identifying objects and faces; large language models like ChatGPT engaging in remarkably human-like conversation.

AI now has the potential to impact everything from the hiring process and loan applications to cybersecurity and financial crime. It can be leveraged for tremendous good, but carries heavy risks if misused. History shows that technological advancements almost always end up having both positive and negative consequences on societies.

Yavuz Sultan Selim Bridge
At 2,164 feet tall, its pylons are the tallest in the world, and its center span of 3,559 feet makes it the world's widest suspension bridge.

Private sector self-regulation is insufficient

As AI rapidly evolves in capability, we need standards and oversight to ensure it develops responsibly. Tech companies, by nature, will choose the path of least resistance in order to maximize profit. Without strong, consistently enforced regulation,

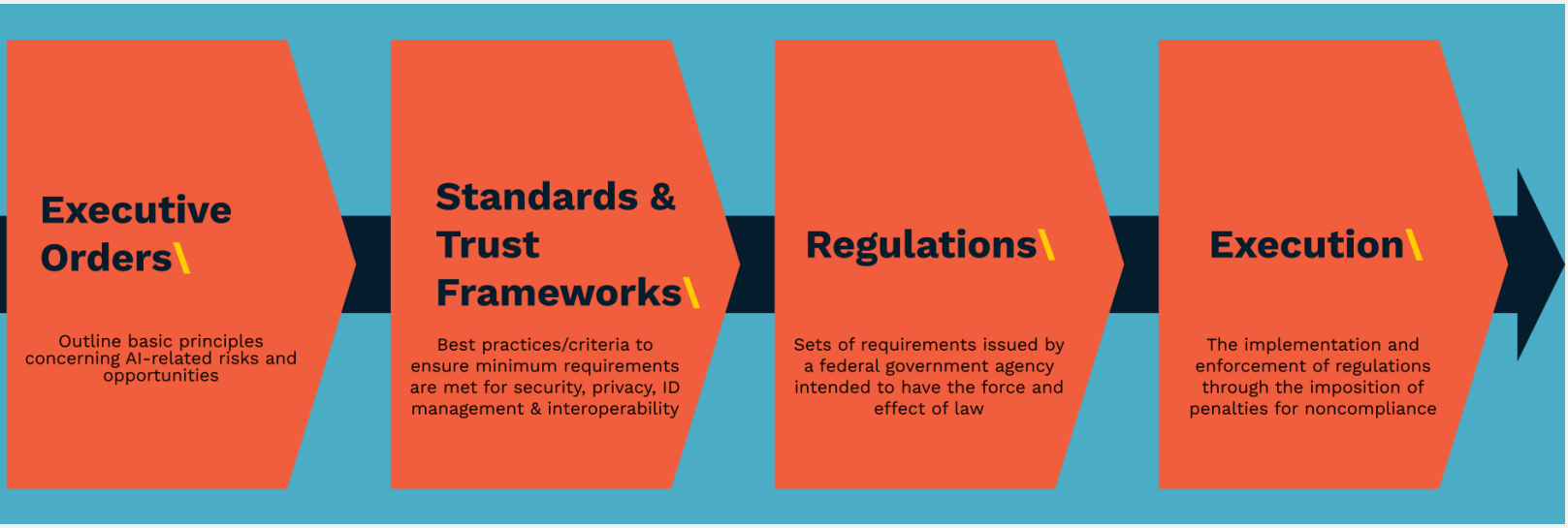
we run the risk of AI progressing without ethics and accountability.

Government involvement in eliminating AI bias is essential to ensure fairness and societal trust in the technology. As machine learning systems increasingly impact critical sectors like healthcare, finance, and law enforcement, mandating unbiased outcomes—and holding accountable those who fail to properly develop bias-free solutions—becomes imperative.

Effective regulation ensures that AI technologies serve humanity's best interests while preventing the propagation of inequitable practices.

Means of management

Governments around the world are using a number of approaches to put pressure on private companies to eliminate bias from AI. Ideally, they deploy a multi-pronged strategy that includes the measures below.



Ideas that carry weight
Principles alone cannot be enforced—they must be codified into regulations to truly compel compliance and enable oversight.

- **Legislation**: Some governments have passed or are considering passing legislation that would regulate the development and use of AI. For example, the European

Union's proposed Artificial Intelligence Act would require companies to assess the fairness of their AI systems and take steps to mitigate bias.

- **Regulation**: Governments are also issuing regulations that govern specific applications of AI. The US Department of Housing and Urban Development, for instance, has issued regulations that prohibit the use of AI in housing discrimination. An increasing number of global regulators are asking entities to carefully select if they use human-in-the-loop, human-out-the-loop, or human-over-the-loop AI systems to ensure that the models' decisions have the right level of human involvement depending on the use case.

- **Trust frameworks**: The Australian Trusted Digital Identity Framework (TDIF) and UK Digital Identity Assurance Trust Framework (DIATF) are national initiatives that outline requirements and standards for digital identity services to securely verify identities online. They aim to enable trusted digital identity transactions between government, businesses, and individuals through certified identity providers and authenticated credentials, and are all-encompassing by covering aspects of data privacy, security and business continuity.

- **Guidance**: Additionally, governments are issuing guidance to businesses on how to develop and use AI in an ethical way. In one such example, the US National Institute of Standards and Technology (NIST) has published guidelines on managing bias in AI. In addition, the Federal Trade Commission has published guidelines on biometric information and Section 5 of the Federal Trade Commission Act. This approach applies a broad-brush approach to set the tone of expectations among the business community in the development and rollout of AI systems.

- **Public awareness**: Governments and quasi-government organizations are running campaigns to raise public

awareness of the issue of bias in AI. This can help to encourage businesses to adopt bias-mitigating measures when developing AI systems.

A look around the world

Globally, government entities in various jurisdictions are proactively addressing AI bias through strategic measures. What follows are some specific examples of what governments are doing to address bias in AI.

The **European Union** is developing a new AI regulation that would require companies to assess the fairness of their machine learning systems and take steps to mitigate bias. The regulation would also prohibit the use of AI for certain applications, such as social scoring and mass surveillance.

The **United Kingdom** has published a set of ethical guidelines for the development and use of AI. The guidelines call for AI systems to be developed in a way that is fair, transparent, and accountable.

Australia has had several governmental agencies, including the Australian Human Rights Commission and the Department of Industry, Science, Energy, and Resources, collaborate on developing AI ethics guidelines to ensure machine learning systems are built and deployed in ways that respect human rights, fairness, and transparency.

New Zealand established the Algorithm Charter for Aotearoa New Zealand, which aims to promote ethical and transparent government use of algorithms. This charter emphasized fairness, transparency, and accountability in the use of AI.

In the **United States**, in January 2021, the White House issued guidelines stating that automated systems should be designed and used equitably, with proactive measures to promote fairness and prevent unjustified discrimination based on protected characteristics.

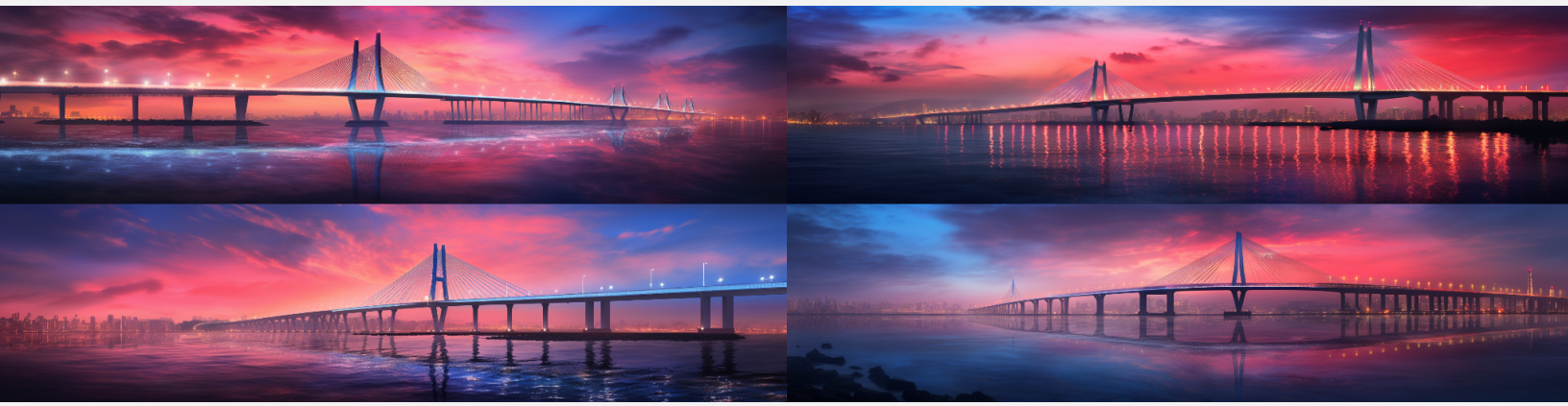
In April 2023, joint guidelines were issued by the Consumer Financial Protection Bureau, the Department of Justice's Civil Rights Division, the Equal Employment Opportunity Commission, and the Federal Trade Commission on enforcement efforts against discrimination and bias in automated systems.

More recently, in October 2023, the White House unveiled the Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence, which was followed swiftly by the Office of Management and Budget (OMB) release of Implementation Guidance in response. These guidelines will directly impact major tech companies as well as smaller tech vendors that serve the Federal system, ultimately percolating into the broader market.

Bypassing Congress, the Executive Order focuses on accountability and responsible innovation in AI systems. Two major aspects include mandating the watermarking of AI-generated content and mitigating AI-driven discrimination.

The watermarking mandate, to be spearheaded by the Commerce Department, will require labeling of all AI-generated audio, visual, and text content. This enables consumers to discern what is human-created versus machine-created, combating deceptive deepfakes. It also promotes transparency and accountability in the AI industry to disclose the data used to train models.

The order also tackles AI-driven discrimination by providing guidance to minimize biased outcomes. It emphasizes inclusivity and fairness in AI applications. This aligns with the concept of



Bandra-Worli Sea Link
This 5.6 km long, 8-lane wide cable-stayed bridge is the longest sea bridge in India.

Zero-Bias AI™ discussed in Part 1 of this report, where engineers adhere to an ethical framework for transparency, consent, data security, and compliance.

The executive order signifies how the White House is taking a proactive approach to managing the risks and opportunities of AI, by directing meaningful regulations to be enacted on key areas like content labeling and non-discrimination. Oversight and responsibility are critical as AI capabilities rapidly advance.

The order also marks a pivotal moment—we have reached a technological threshold where meaningful oversight and governance are vital. Executive orders alone are insufficient; they must be codified into laws to have enforceable impact.

It takes an entire society

There exists a shared responsibility among governments, private enterprises, and individuals to shape a future where artificial intelligence is a force for good for all. As discussed above, governments must craft robust strategies that prioritize ethical AI deployment to safeguard against biases and ensure maximum transparency.

Meanwhile, private businesses retain the duty to design and develop AI technologies that adhere to these principles, meshing their drive for innovation with accountability. And individual humans, as the ultimate protectors of their own interests, must be strong advocates for unbiased AI and take it upon themselves to understand the implications of putting machine learning algorithms in charge of making critical decisions.

About the imagery

IDVerse's dedication to pushing boundaries in the realm of identity verification with generative AI has led us to explore the applications of this technology in other contexts. Hence, we made the choice to incorporate the artificially produced visuals in this document.

For this report, we consider bridges as an embodiment of standards or frameworks, providing a powerful metaphor for understanding the critical role that regulations play in the functioning of various systems. Just as bridges are engineered to connect two distinct points, regulations serve as the vital links that create a structured pathway, ensuring the smooth flow of activities and transactions while mitigating potential risks and hazards that may arise along the way.

About the author

Terry Brenner *is Head of Legal, Risk & Compliance, Americas, for IDVerse. Previously he has served in executive office and general counsel roles, in both start-up and mature businesses, across a range of diverse industry sectors. His focus at IDVerse is to lay the path for the successful integration of IDVerse's remote ID verification technology into the Americas market, heeding to the sensitivities around data and privacy protection. From a commercial perspective, he drives towards supporting disciplined growth of the business whilst reinforcing the ethics and compliance mission of IDVerse to be the industry benchmark from a compliance perspective and to build trust in the brand.*

About IDVerse

IDVerse, an OCR Labs company, is the leading automated identity verification platform to onboard and re-authenticate trusted users at scale.

What sets us apart? Our commitment to Zero Bias AI™ means that we are pioneering the use of machine learning to protect against discrimination on the basis of ethnicity, age, and gender. We build software capable of authenticating tens of thousands of ID document types and verifying the liveness of billions of real people without manual human intervention—all underpinned by generative AI that achieves maximum inclusion and fairness.

IDVerse can recognize over 16,000 ID types in 142 languages from more than 230 countries and territories. The world's leading companies like Amex, HSBC, and Hertz trust us to help their users prove their identity in seconds.

The IDVerse solution has been tested and certified to meet the most stringent standards in the industry, including NIST, ISO, iBeta, and algorithmic Zero Bias AI™ specifications.

Want to learn more? Book a demo today, or get in touch with us at hello@idverse.com.