WORLD GOVERNMENTS SUMMIT 2025

in collaboration with KPMG

REPORT

The Future of AI Governance

The UAE Charter and Global Perspectives

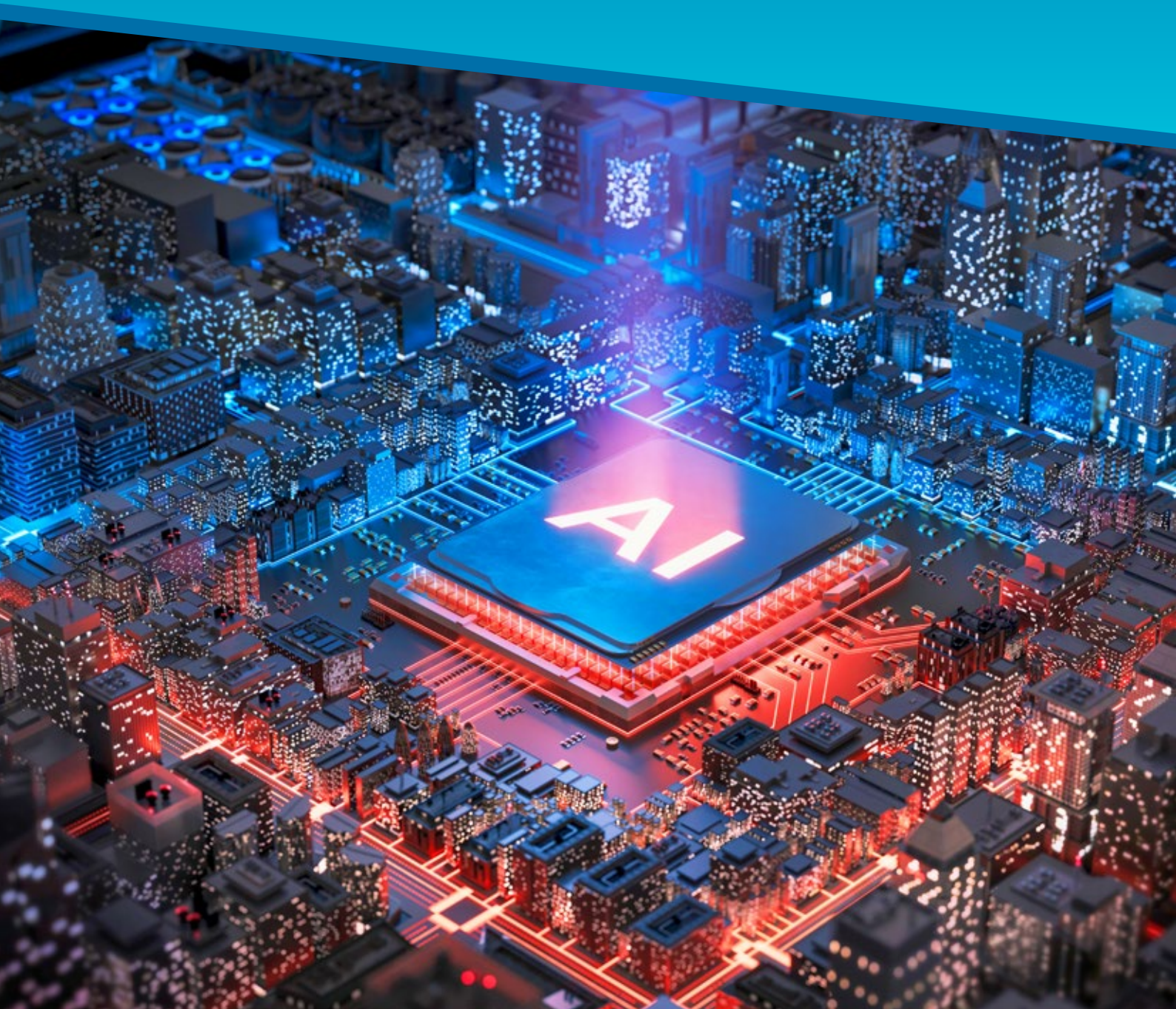


Table of Contents

Topics

Foreword

We recognize that a clear, actionable set of AI principles forms the cornerstone of ethical and responsible AI development. These principles are not only essential for building public trust and ensuring organizational accountability, but also for fostering inclusive innovation that benefits citizens, businesses, and governments alike.

As global regulatory frameworks evolved, such as the EU AI Act passed in 2024, grounded in the European Commission's ethical guidelines for trustworthy AI, principles-based governance has emerged as the foundational approach to AI oversight. The UAE has demonstrated regional and global leadership through its AI Strategy 2031 and the release of the UAE AI Charter for the development and use of Artificial Intelligence, in July 2024, which articulates twelve key principles to ensure AI is deployed safely, equitably, and transparently.

This whitepaper offers a detailed interpretation of each of the 12 UAE AI Charter principles, actionable recommendations for implementation across the AI lifecycle, mapped to KPMG's Trusted AI Framework, practical insights to support AI governance, risk management, and regulatory alignment, and a blueprint for building resilient, human-centric AI systems in alignment with the UAE's national priorities.

The UAE Charter places particular emphasis on human oversight, inclusivity, safety, and legal compliance—values that resonate with global AI ethics standards like those outlined by OECD, UNESCO, and the EU. As AI regulation becomes more stringent, organizations that proactively align with these principles will be better positioned to lead responsibly, mitigate risks, and capture the full potential of AI innovation.

To move beyond aspirational intent, organizations must embed these principles into operational reality. This means evolving existing governance models to support the distinct requirements of AI—such as data provenance tracking, model accountability, explainability, bias audits, and human oversight. Governance frameworks must shift from static policies to adaptive controls that align with the fast-evolving AI lifecycle.

Embedding the UAE AI Charter into enterprise governance also provides a strategic advantage. It signals readiness for future compliance, enables risk-aware innovation, and ensures that AI deployments are not only lawful but also aligned with public expectations and societal values. Organizations that operationalize these principles early will be better equipped to manage ethical dilemmas, respond to regulatory inquiries, and build lasting trust with users, regulators, and the wider community.

Proactively implementing these principles not only ensures regulatory readiness but also delivers clear business value. Organizations that embed responsible AI practices early can accelerate innovation with confidence, reduce compliance costs, and enhance their reputation as trustworthy, forward-thinking leaders. By building AI systems that are transparent, inclusive, and human-centric, businesses can unlock new opportunities, gain stakeholder trust, and differentiate themselves in an increasingly AI-driven economy.

Across the globe, jurisdictions such as the European Union, Canada, the United States, and Singapore are moving swiftly to codify AI ethics into binding legislation and operational frameworks. This signals a global shift where AI governance will no longer be optional—but a core component of digital competitiveness and enterprise resilience.

The UAE Charter: The 12 AI Principles

1. Strengthening Human-Machine Ties:

The UAE aims to enhance the harmonious and beneficial relationship between AI and humans, ensuring that all AI developments prioritize human well-being and progress.

2. Safety:

The UAE places great importance on safety, ensuring that all AI systems comply with the highest safety standards. The country encourages modifying or removing systems that pose risks.

3. Algorithmic Bias:

The UAE aims to address the challenges posed by AI algorithms regarding algorithmic bias, contributing to a fair and equitable environment for all community members. This promotes responsible development of AI technologies, making them inclusive and accessible to everyone, supporting diversity, and respecting individual differences. It ensures equal technological benefits and improves quality of life without exclusion or discrimination.

4. Data Privacy:

In line with the UAE's stance on privacy rights, while data is essential for AI development, supporting and promoting innovation in AI, the privacy of community members remains a top priority.

5. Transparency:

The UAE seeks to create a clear understanding of AI and how systems operate and make decisions, which helps build trust, enhance responsibility, and promote accountability in the use of these technologies.

6. Human Oversight:

The Charter emphasizes the irreplaceable value of human judgment and human oversight over AI, aligning with ethical values and social standards to correct any errors or biases that may arise.

7. Governance and Accountability:

The UAE adopts a responsible and proactive stance, emphasizing the importance of governance and accountability in AI to ensure the technology is used ethically and transparently.

8. Technological Excellence:

AI should be a beacon of innovation, reflecting the UAE's vision of digital, technological, and scientific excellence. The UAE seeks global leadership by adopting technological excellence in AI to drive innovation, enhance competitiveness, and improve quality of life through innovative and effective solutions to complex challenges, contributing to sustainable progress benefiting society as a whole.

9. Human Commitment:

Human commitment in AI reflects the spirit of the UAE, essential for ensuring that the development of this technology serves the public good. It focuses on enhancing human well-being and protecting fundamental rights, emphasizing the importance of placing human values at the heart of technological innovation to ensure a positive and lasting impact on society.

10. Peaceful Coexistence with AI:

Peaceful coexistence with AI is crucial to ensure technology enhances the well-being and progress of our communities without compromising human security or fundamental rights.

11. Promoting AI Awareness for an Inclusive Future:

It is essential to create an inclusive future that ensures everyone can benefit from AI advancements, guaranteeing equitable access to this technology and its advantages for all segments of society.

12. Commitment to Treaties and Applicable Laws:

The UAE emphasizes the importance of complying with international treaties and local laws in the development and use of AI.

KPMG's Trusted AI Framework

As artificial intelligence becomes increasingly integral to critical decisions and everyday operations, KPMG developed its Trusted AI Framework to help organizations navigate this evolving landscape. The framework brings structure, accountability, and clarity to the AI lifecycle, ensuring that AI systems are ethical, transparent, and aligned with human values from strategy to deployment.

Built on KPMG's global experience across industries, the framework is founded on ten core principles. These principles include fairness and transparency, which ensure AI systems are inclusive and understandable; explainability and accountability, which foster human oversight and responsibility; and privacy, security, and safety, which protect both individuals and systems. Additionally, the framework emphasizes data integrity and reliability for consistent AI performance, as well as sustainability to ensure AI advancements contribute to broader social and environmental goals.

The UAE AI Charter reflects a similar commitment to responsible AI development, expressing a national vision through twelve guiding principles that align closely with those in KPMG's Trusted AI Framework. The table below illustrates how each UAE AI principle maps to one or more of KPMG's Trusted AI principles:

	UAE AI Principle	Aligned KPMG Global Trusted AI Principle(s)
1.	Strengthening Human-Machine Ties	Explainability, Fairness, Accountability
2.	Safety	Safety, Reliability, Security
3.	Algorithmic Bias	Fairness, Transparency, Data Integrity
4.	Data Privacy	Privacy, Data Integrity
5.	Transparency	Transparency, Explainability
6.	Human Oversight	Accountability, Explainability
7.	Governance and Accountability	Accountability, Transparency, Data Integrity
8.	Technological Excellence	Reliability, Sustainability
9.	Human Commitment	Fairness, Sustainability, Accountability
10.	Peaceful Coexistence with AI	Safety, Security, Fairness
11.	Promoting AI Awareness for an Inclusive Future	Fairness, Explainability
12.	Commitment to Treaties and Applicable Laws	Accountability, Privacy, Data Integrity

This close alignment between the UAE AI Charter and KPMG's Trusted AI Framework provides a strong foundation for action. The Trusted AI principles have already been operationalized through defined methodologies across the AI lifecycle—spanning strategy and design, data enablement, model development, testing and evaluation, and deployment and monitoring. Building on this proven foundation, the same structured approach has been applied in this whitepaper to the UAE's twelve AI principles. For each, practical steps are outlined to help organizations embed ethical, human-centric AI practices and turn principles into tangible outcomes.

Principle 1

Strengthening Human-Machine Ties

The UAE aims to enhance the harmonious and beneficial relationship between AI and humans, ensuring that all AI developments prioritize human well-being and progress.

Understanding the Principle in Real-World Terms

This principle aims to ensure that AI systems enhance and augment human capabilities, empowering human beings to exceed their potential by creating smarter, more inclusive solutions. AI should align with ethical principles, respecting human dignity, rights, and values. Ultimately, the UAE aims to foster an environment where humans and AI collaborate to improve quality of life, boost productivity, and drive societal progress.

Real-world examples:

- **Healthcare AI:** AI systems used in healthcare to assist doctors in diagnosing diseases more accurately and efficiently, ultimately improving patient outcomes.
- **Smart Cities:** AI-driven technologies integrated into urban planning to improve infrastructure, optimize traffic flow, and enhance the quality of life for residents.

Embedding This Principle into AI Governance

To strengthen human-machine ties, focus on developing AI systems that augment human capabilities, enhance well-being, and drive positive outcomes for employees, customers, and society. Ensure that your AI initiatives align with best practices and thoughtfully consider their broader impact on human values, dignity, and rights, while also reflecting the cultural and societal values of the UAE throughout every stage of development and implementation. Consider incorporating human-in-the-loop decision-making to further reinforce the human-AI relationship, ensuring meaningful oversight, trust, and accountability.

Best Practices and Methodologies

Strategy and Design

- **Human-Centric Design:** Design AI systems to augment human capabilities by continuously gathering diverse feedback and using it to refine and enhance AI's impact.
- **Ethical AI Goals:** Set clear ethical guidelines for AI development that prioritize human well-being and address potential negative impacts, ensuring alignment with both global standards and UAE's cultural values.
- **Human-in-the-Loop Integration:** Consider embedding human-in-the-loop mechanisms early to strengthen decision-making, ensure accountability and align AI systems with core human values.

Data Enablement

- **Data Sensitivity:** Build trust by ensuring data collection and processing respects human privacy and dignity, ensuring the responsible use of personal and sensitive data.
- **Well-being Metrics:** Consider factors such as well-being, safety, and user experience when evaluating the data used for training AI models.
- **Diverse Data Representation:** Use datasets that reflect diverse human experiences, ensuring AI systems can serve the broad spectrum of societal needs.

Model Development

- **Human-AI Collaboration Features:** Develop AI systems that enhance human capabilities by offering insights, and providing support, without replacing human decision-making.
- **Transparent Algorithms:** Build models that allow humans to easily understand, trust, and collaborate with AI systems. Transparency helps ensure human oversight is maintained.
- **Bias Reduction:** Ensure AI models are free from biases that may harm human progress, including ensuring equitable treatment across diverse groups and preserving societal values.

Testing and Evaluation

- **Human Impact Assessment:** Assess the potential positive and negative impacts of AI systems developed on human well-being, ensuring the outcomes are aligned with the intended benefits.
- **User Feedback:** Incorporate feedback from users to fine-tune AI systems, ensuring they are relevant to human needs and progress.

Deployment and Monitoring

- **Continuous Collaboration:** Maintain active collaboration with human users and stakeholders to ensure that deployed AI systems remain beneficial and enhance human progress.
- **Monitor AI for Human Impact:** Track the long-term effects of AI on society, ensuring AI systems continue to prioritize and enhance human well-being.
- **Adaptation to Human Needs:** Continuously adapt AI technologies to meet the evolving needs and values of human users, especially as societal contexts change.

Extending the Principle to Agentic AI Systems

Agentic AI systems must be designed to complement, not replace, human roles. They should enhance human decision-making and productivity through contextual awareness and feedback mechanisms. Ensuring intuitive human interaction and transparency will help preserve trust. Organizations must prioritize user experience in agent-AI interfaces. Emotional and cognitive impact on users should be monitored and improved over time.

Key Tools, Techniques and Further Reading

Tools and Techniques:

- Human-Centered AI Design Frameworks
- Human-AI Collaboration Toolkits (e.g. Microsoft Copilot, Salesforce Einstein)
- UX Research and Cognitive Load Testing tools (e.g. Optimal Workshop)
- KPMG Trusted AI Framework
- KPMG Trusted AI Risk and Control Matrix (RCM)

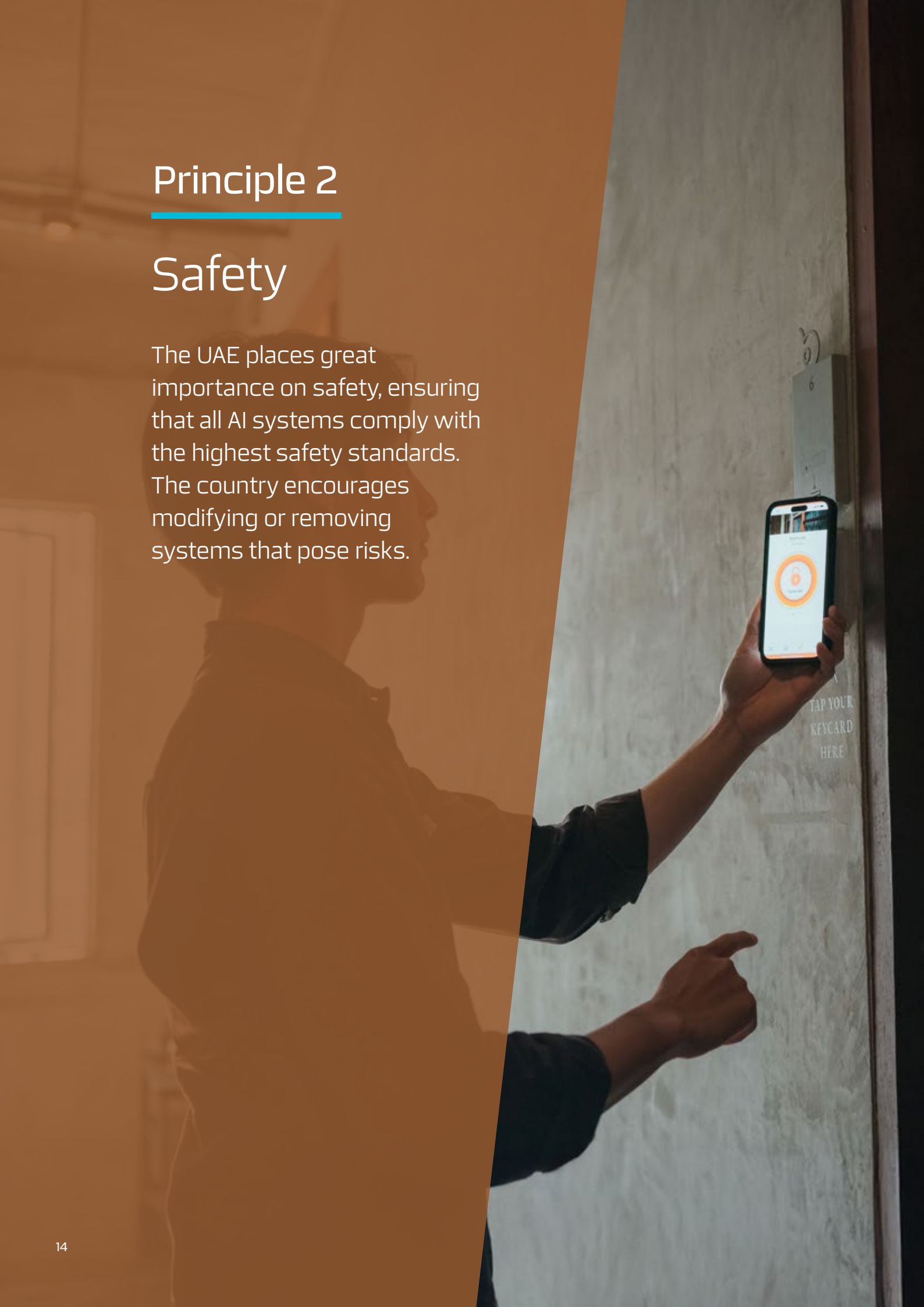
Further Reading:

- Stanford HAI: Human-AI Collaboration Studies
- Harvard Berkman Klein Center: Ethics of Augmentation
- Microsoft: The Future Computed – AI and Human Values

Principle 2

Safety

The UAE places great importance on safety, ensuring that all AI systems comply with the highest safety standards. The country encourages modifying or removing systems that pose risks.



Understanding the Principle in Real-World Terms

AI safety refers to ensuring that AI systems function as intended, without causing harm to individuals, businesses, or society. This includes technical robustness, risk mitigation, and incorporating fail-safes to prevent or address unintended consequences or failures. The EU AI Act exemplifies this approach by mandating stringent safety standards for high-risk AI applications. Prioritizing safety is essential to minimizing risks and maintaining both operational continuity and public trust in the UAE.

Real-world examples:

- **Autonomous Vehicles:** AI-driven cars failing to recognize pedestrians in low visibility conditions, leading to accidents and regulatory scrutiny.
- **Healthcare AI:** Diagnostic AI misinterpreting medical images, leading to incorrect treatments and potential liability risks.

Embedding This Principle into AI Governance

Ensuring AI safety requires a structured approach—from risk assessments to continuous testing and fail-safe mechanisms. Tools like KPMG's Trusted AI Risk Framework support this process by offering a structured methodology to identify, assess, and mitigate AI-related risks, including those tied to safety, in alignment with standards such as ISO 42001 and the EU AI Act. Combined with a strong AI governance framework, these tools help ensure safety measures remain effective, transparent, and aligned with both local and global best practices. By embedding safety best practices into every stage of development, organizations can enhance reliability, maintain compliance, and build trust.

Best Practices and Methodologies

Strategy and Design

- **Safety Goals and Metrics:** Establish clear safety goals and metrics for AI initiatives, focusing on reliability, resilience, transparency, and security. Tools like KPMG's AI metrics can measure performance and ensure ethical alignment.
- **Risk Identification:** Define safety risks associated with the AI systems and establish protocols for risk mitigation.
- **Stakeholder Consultation:** Engage regulators, industry experts, and end-users to anticipate safety concerns before development.
- **Fail-Safe Design:** Ensure that the design of AI systems has clear override mechanisms to prevent harm in case of failure. Incorporate fallback mechanisms, monitoring, and human-in-the-loop.

Data Enablement

- **Data Integrity Checks:** Validate training data for accuracy, completeness, and consistency to prevent AI failures.
- **Bias and Anomaly Detection:** Identify biases that could lead to unsafe AI behavior, such as misclassification in healthcare or autonomous systems.
- **Simulation and Stress Testing Data:** Train AI models on varied scenarios, including edge cases, to ensure robustness in real-world applications.

Model Development

- **Safety-Conscious Algorithms:** Implement algorithms that prioritize safety, incorporating guardrails and constraints to prevent harmful decisions.
- **Fail-Safe Mechanisms:** Embed fail-safe mechanisms like human-in-the-loop and logging in the final design.
- **Explainability and Transparency:** Ensure that AI decisions can be understood and audited to identify potential safety risks before deployment.

Testing and Evaluation

- **Adversarial Testing:** Identify weak points in the AI system by testing against potential failure points and establish corrective measures before deployment.
- **Testing for Edge Cases:** Evaluate AI performance under extreme conditions (e.g. low visibility for self-driving cars or unpredictable market fluctuations in finance).
- **Safety Benchmarking:** Define and measure safety performance against industry standards.

Deployment and Monitoring

- **Continuous Safety Monitoring:** Regularly audit AI systems post-deployment to detect anomalies or failures.
- **Incident Response Plans:** Establish clear escalation protocols for AI malfunctions to ensure quick remediation.
- **Regulatory Compliance Reporting:** Document and communicate safety measures within the organization to demonstrate adherence to safety standards.

Extending the Principle to Agentic AI Systems

Agentic AI introduces dynamic decision-making, which requires real-time risk detection and mitigation capabilities. Safety protocols must be embedded not just in code, but also in how agents interact with systems and people. Fail-safes and escalation paths to human supervisors are essential. Simulation testing for adversarial or unintended agent behavior must be prioritized. Organizations should track agent actions to ensure accountability.

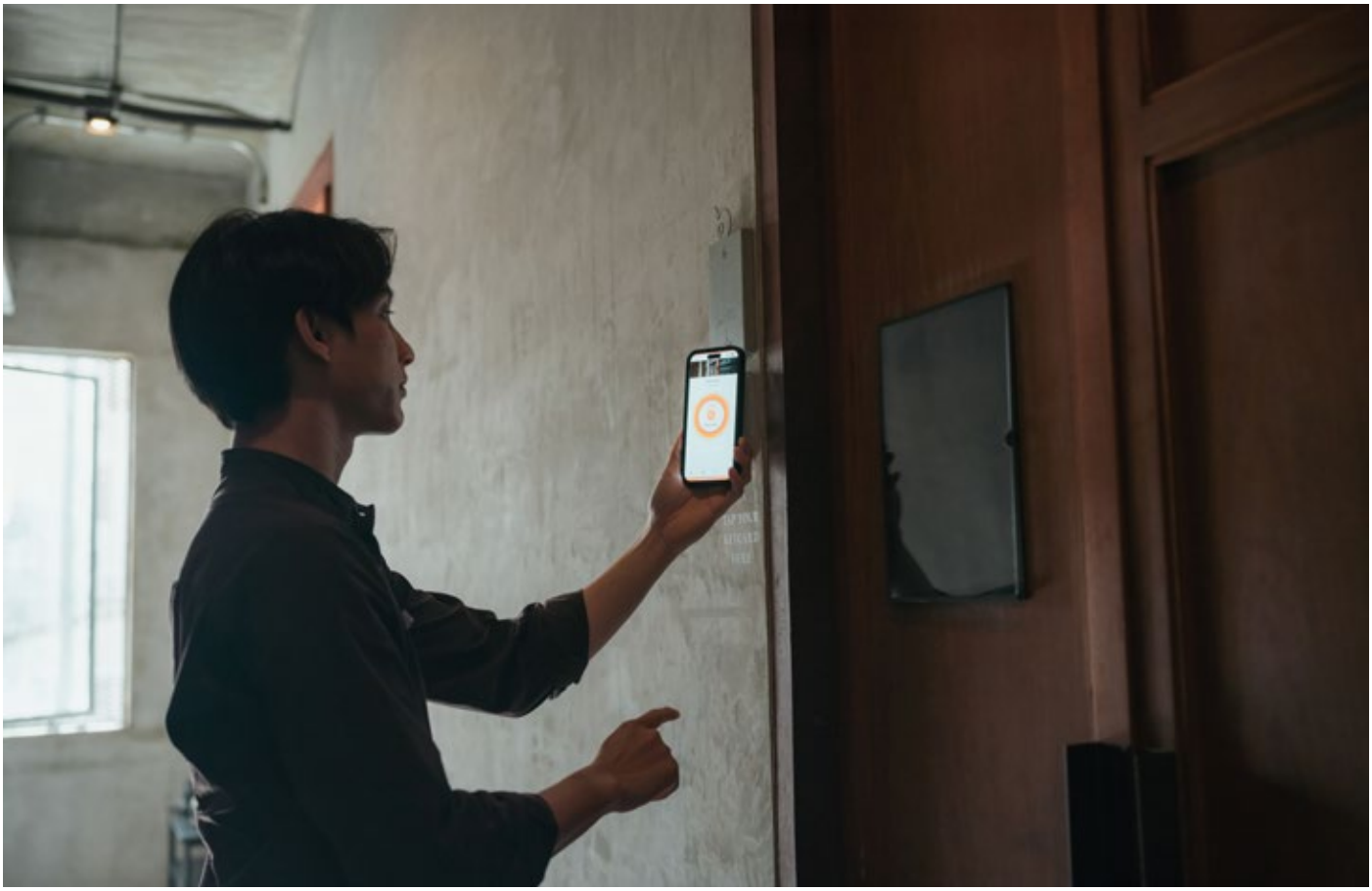
Key Tools, Techniques and Further Reading

Tools and Techniques:

- Adversarial Testing Frameworks (e.g. CleverHans, Foolbox)
- Formal Verification Tools (e.g. TLA+, Z3)
- Bayesian Networks, Monte Carlo Dropout
- Red Teaming and Simulation Labs
- KPMG Trusted AI Framework
- KPMG Trusted AI Risk and Control Matrix (RCM)

Further Reading:

- NIST AI Risk Management Framework
- OpenAI's System Safety Practices
- EU AI Act: Safety Provisions for High-Risk Systems



Principle 3

Algorithmic Bias

The UAE aims to address the challenges posed by AI algorithms regarding algorithmic bias, contributing to a fair and equitable environment for all community members. This promotes responsible development of AI technologies, making them inclusive and accessible to everyone, supporting diversity, and respecting individual differences. It ensures equal technological benefits and improves quality of life without exclusion or discrimination.

Understanding the Principle in Real-World Terms

In practice, algorithmic bias occurs when AI systems make decisions that unintentionally favor or disadvantage certain groups based on factors like gender, ethnicity, age, or socioeconomic status. This can stem from biased training data, flawed model assumptions, or a lack of diverse representation in development. Addressing bias is critical to building trust, ensuring fairness, and mitigating financial, legal, and reputational risks.

Real-world examples:

- **Hiring Systems:** AI systems rejecting female candidates based on biased training data derived from male-dominated industries, exposing the company to discrimination claims or regulatory penalties.
- **Loan Approvals:** AI-based credit systems rejecting loan applicants based on historical discriminatory practices, potentially violating fair lending laws.

Embedding This Principle into AI Governance

To effectively address algorithmic bias in your AI systems, embed fairness into every phase of development. This includes making proactive design choices, using representative and balanced data, and conducting continuous testing to ensure equitable outcomes. Effective AI governance, supported by a strong framework, should be woven into each stage to ensure accountability, transparency, and compliance with ethical and local regulatory standards. By doing so, organizations can translate the principle of fairness into actionable, impactful steps that drive responsible AI development that aligns with the UAE's requirements.

Best Practices and Methodologies

Strategy and Design

- **Define Fairness Objectives:** Clearly outline fairness goals with metrics for each AI initiative, identifying potential biases and ensuring that the needs of diverse stakeholder groups are represented. This should be guided and aligned with your organization's AI governance framework.
- **Inclusive Design:** Involve diverse teams—ranging from ethicists to affected community members—during the design phase to identify and address possible sources of bias before they emerge.
- **Adopt Explainable AI (XAI):** Ensure transparency in how AI models make decisions by implementing explainable AI frameworks that allow stakeholders to understand and audit AI outputs.

Data Enablement

- **Ensure Representativeness:** Make sure datasets accurately represent diverse demographics (age, gender, ethnicity, socioeconomic status, etc.), particularly in areas like hiring, healthcare, and finance.
- **Bias Audits:** Regularly perform audits on training datasets to identify and mitigate historical and societal biases, including direct biases, proxy biases, sampling biases, and measurement biases.
- **Data Labeling:** Improve the quality and accuracy of data labeling to prevent subjective or biased annotations. Ensure that data labels reflect the diversity of the population being represented.

Model Development

- **Bias Detection During Model Design:** Screen AI models for potential proxy biases—variables like zip codes or income levels that could inadvertently reflect sensitive attributes such as race or gender.
- **Fairness-Aware Algorithms:** Use fairness-enhancing algorithms that can identify and correct biases by adjusting for imbalances in the model's predictions or outcomes.
- **Document Model Choices:** Document the rationale behind key decisions made during model development, including the selection of features and algorithms, to ensure transparency and accountability.

Testing and Evaluation

- **Threshold Setting:** Before deployment, define acceptable fairness thresholds that align with the fairness goals and metrics set at the ideation stage.
- **Impact Testing:** Test the fully trained models for bias against the fairness thresholds and evaluate how the AI system's outcomes differ across demographics and adjust as necessary to ensure equitable results.

Deployment and Monitoring

- **Ongoing Monitoring:** Continuously monitor the performance of AI systems post-deployment to ensure fairness and alignment with the governance framework. Ensure the system adapts to evolving societal norms and expectations.

- **Feedback Mechanisms:** Establish mechanisms that allow users and stakeholders to report biased outcomes or issues.
- **Periodic Reporting:** Regularly publish bias impact testing reports, allowing both internal and external stakeholders to hold the organization accountable.

Extending the Principle to Agentic AI Systems

Bias in agentic AI can compound through autonomous loops and adaptive behavior. Fairness must be evaluated continuously as agents evolve their decision rules. Diverse training environments and dynamic bias audits are necessary. Agents must be restricted from reinforcing systemic biases via feedback loops. Interventions must be logged and governed via ethical review boards.

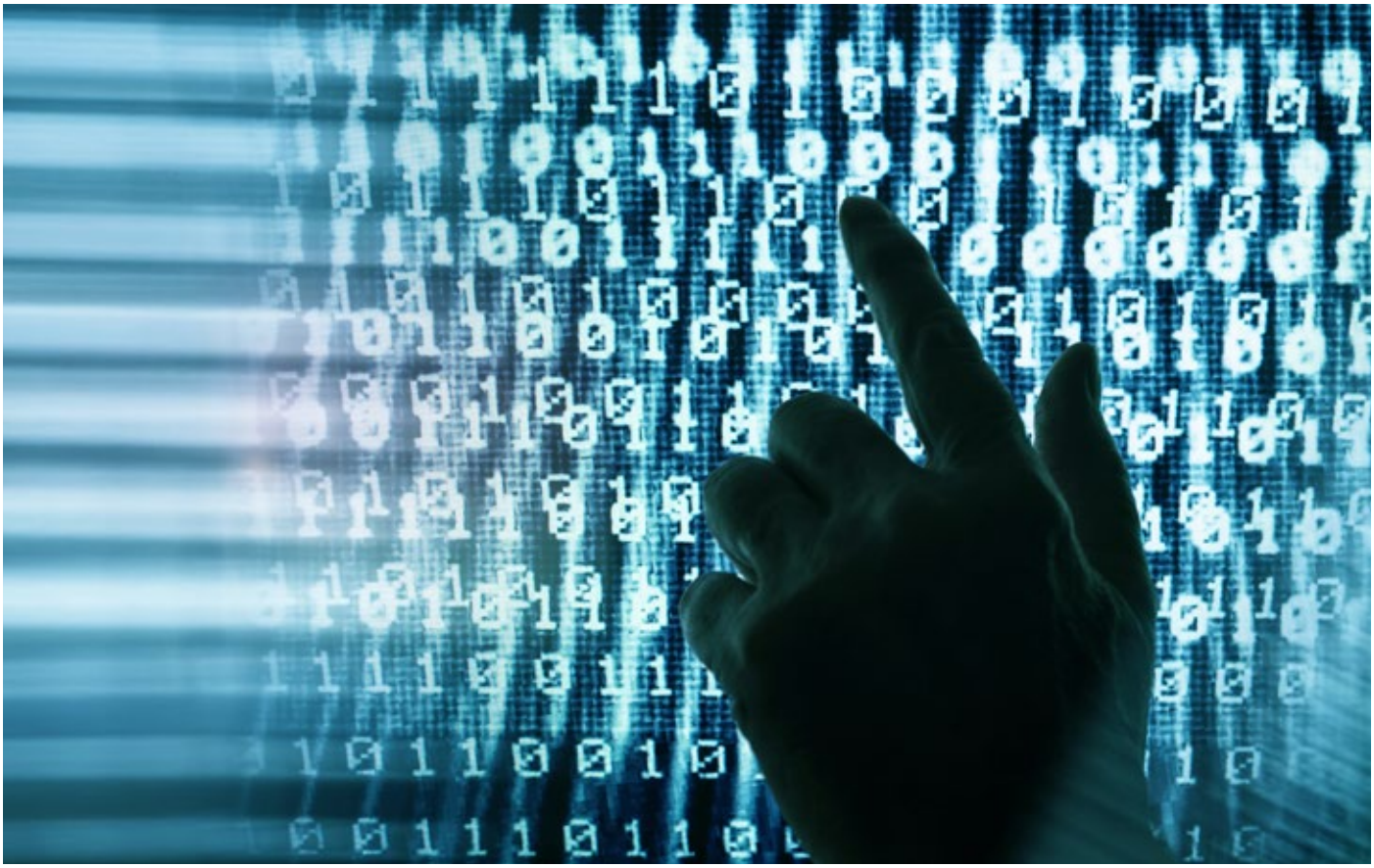
Key Tools, Techniques and Further Reading

Tools and Techniques:

- IBM Fairness 360
- Microsoft Fairlearn
- Google What-If Tool
- SHAP, LIME for explainable fairness analysis
- KPMG Trusted AI Framework
- KPMG Trusted AI Risk and Control Matrix (RCM)

Further Reading:

- World Economic Forum: Fairness Toolkit
- Partnership on AI: Managing Bias in AI
- AI Now Institute: Algorithmic Accountability Report



Principle 4

Data Privacy

In line with the UAE's stance on privacy rights, while data is essential for AI development, supporting and promoting innovation in AI, the privacy of community members remains a top priority.

Understanding the Principle in Real-World Terms

Data privacy in AI refers to the protection of personal and sensitive information throughout the AI lifecycle — from data collection and storage to model training and deployment. As AI models depend on vast datasets, organizations must ensure that individuals' data remains confidential and is used responsibly, with transparency and consent. Failure to safeguard privacy can result in regulatory penalties, reputational damage, and erosion of public trust, all of which can undermine long-term innovation efforts.

Real-world examples:

- **Retail AI:** Algorithms using customer purchase history without explicit consent for personalized marketing, resulting in customer complaints and data protection violations.
- **Healthcare AI:** Medical AI models trained on patient data shared between hospitals and third parties without proper consent, leading to confidentiality breaches and legal action.

Embedding This Principle into AI Governance

Organizations should build data privacy into every stage of the AI lifecycle. This means limiting data collection to what's necessary, obtaining clear consent, ensuring secure storage, and embedding privacy-by-design in system architecture. Accountability structures should be clear, with defined ownership for privacy oversight. Factors such as secure data handling, explicit consent, and safeguards for personal data are also emphasized in the EU AI Act — reinforcing the need for organizations to treat privacy as a core design principle, not an afterthought. Doing so reduces regulatory risk, builds trust, and positions the organization as a responsible leader.

Best Practices and Methodologies

Strategy and Design

- **Set Privacy Objectives:** Define privacy standards for each AI system, aligned with legal and organizational guidelines, and define metrics to monitor compliance throughout the AI lifecycle.
- **Privacy Impact Assessments (PIAs):** Conduct PIAs during the early planning stages to anticipate and mitigate privacy risks.
- **Design for Privacy and Confidentiality:** Incorporate design principles such as data minimization, anonymization, and opt-out privileges into system architecture and user experiences.
- **Third-Party Model Assessment:** When using third-party models, evaluate the vendor's data privacy and confidentiality protocols to ensure they meet required standards.

Data Enablement

- **Data Minimization and Anonymization:** Limit the collection of personal data and anonymize datasets wherever possible.
- **Consent and Usage Clarity:** Ensure all data is collected with informed, explicit consent, maintain an audit trail, and provide clear data usage terms and opt-out processes.
- **Access and Classification Controls:** Restrict data access to authorized personnel only, log all activity, and implement tagging taxonomies to flag protected data and restricted attributes.
- **Responsible Aggregation and Third-Party Use:** Protect sensitive information when aggregating sources and verify third-party data complies with privacy and contractual standards.

Model Development

- **Privacy-Preserving Techniques:** Incorporate differential privacy, federated learning, and other techniques that protect individual data while training models.
- **Model Explainability:** Provide clear documentation of how data is used and how models make decisions using that data.

Testing and Evaluation

- **Data Leakage Testing:** Validate that models do not inadvertently leak sensitive data.
- **Privacy Stress Testing:** Simulate potential privacy breach scenarios and verify the effectiveness of safeguards.
- **Risk Thresholds:** Define acceptable risk thresholds related to privacy and ensure models comply before deployment.

Deployment and Monitoring

- **Real-Time Monitoring:** Continuously monitor AI systems for privacy compliance issues or anomalies.
- **Incident Response Protocols:** Establish escalation procedures for potential data breaches or misuse incidents.
- **Privacy Reports:** Regularly publish reports on privacy controls and incidents to promote stakeholder trust.

Extending the Principle to Agentic AI Systems

Agentic AI systems often rely on continuous data ingestion and contextual awareness, raising complex privacy concerns. Consent mechanisms must adapt in real time, and data use must be auditable. Agents should operate on the principle of data minimization and need-to-know basis. Federated learning or edge processing can reduce central data exposure. Privacy risk assessments should be updated with each new agent capability.

Key Tools, Techniques and Further Reading

Tools and Techniques:

- Differential Privacy Libraries (Google DP, OpenDP)
- TensorFlow Privacy
- Data Masking and Anonymization Tools (e.g. ARX, Privitar)
- Consent Management Platforms (e.g. OneTrust)
- KPMG Trusted AI Framework
- KPMG Trusted AI Risk and Control Matrix (RCM)

Further Reading:

- UAE Federal Data Protection Law
- GDPR Guidelines on AI
- NIST Privacy Framework

Principle 5

Transparency

The UAE seeks to create a clear understanding of AI and how systems operate and make decisions, which helps build trust, enhance responsibility, and accountability in the use of these technologies.

Understanding the Principle in Real-World Terms

AI transparency ensures responsible disclosure to stakeholders by providing clear insight into how AI systems function across the entire lifecycle. It ensures that stakeholders understand when they are interacting with AI, how decisions are made, and what data or limitations are involved. Transparent systems support accountability, enable informed oversight, and help build trust by making AI understandable and interpretable to users, regulators, and affected parties. Without transparency, AI can become a "black box," increasing the risk of biased, harmful, or unchallengeable outcomes.

Real-world examples:

- **Loan Approvals:** Customers being denied loans by AI systems without an explanation, leaving them dissatisfied and without recourse.
- **Healthcare AI:** Patients not understanding why an AI system recommended a certain treatment over another, leading to reduced confidence in the system and potential challenges to medical decisions.

Embedding This Principle into AI Governance

To ensure AI transparency, businesses should adopt explainable AI models, maintain clear documentation, and embed responsible disclosure mechanisms across the AI lifecycle. By integrating best practices within a robust AI governance framework, organizations can align with the UAE's transparency principle while fostering trust and confidence in AI-driven decisions.

Best Practices and Methodologies

Strategy and Design

- **Identify Stakeholders:** Document all user groups and define how they will interact with the AI system.
- **Define Transparency Objectives:** Identify what level of explainability is required for different users (e.g. regulators, customers, internal teams).
- **Design Disclosure Mechanisms:** Inform users when they interact with AI, obtain consent for data use, disclose data refresh rates, and communicate known data or model limitations.
- **Design with Explainability:** Build AI systems with explainability in mind, defining success criteria and metrics while considering stakeholder needs.

Data Enablement

- **Data Provenance Tracking:** Maintain records of data sources, ensuring that training data is traceable and auditable.
- **Data Interpretability:** Ensure that the input data used by AI models is understandable and well-documented, so stakeholders can trace how data influences AI decisions.
- **User-Interpretable Data Inputs:** Ensure AI models rely on inputs that can be logically explained, avoiding obscure or opaque variables.

Model Development

- **Explainable AI (XAI):** Use interpretable models where possible or implement techniques that explain how complex models arrive at their decisions.
- **Feature Importance Analysis:** Clearly document key features that influence AI decisions, helping to clarify which data points have the most impact on outcomes.
- **Develop Disclosure Mechanisms:** Integrate responsible disclosure tools and frameworks directly into the model architecture.
- **Decision Logging and Audit Trails:** Maintain logs of AI decision-making processes to enable transparency, facilitate audits and support reviews in case of disputes.

Testing and Evaluation

- **User Testing for Interpretability:** Test AI systems with end users to determine if explanations are understandable and actionable.
- **Fairness and Explainability Metrics:** Establish quantifiable metrics to measure how transparent AI decisions are and adjust where necessary.

Deployment and Monitoring

- **Ongoing Monitoring:** Continuously track AI decisions and update governance frameworks to adapt and ensure models remain interpretable.
- **Publish System Documentation:** Clearly explain model type, intended use, operational conditions, limitations, and data sources.
- **Transparency Reports:** Publish regular reports outlining how AI systems operate, their capabilities and limitations including potential risks and safeguards, to maintain accountability.

Extending the Principle to Agentic AI Systems

With agentic AI, decision chains may span multiple models and interactions. Organizations must ensure agents can explain their choices and provide traceable logic. Visual or narrative explanations should be designed for user comprehension. Transparency logs should include triggers, intent, and reasoning. Role-based explainability may be needed—for users, auditors, and regulators. Periodic reviews must evaluate how well agent behavior aligns with declared purpose.

Key Tools, Techniques and Further Reading

Tools and Techniques:

- SHAP, LIME, ELI5, Skater
- Model Cards, Datasheets for Datasets
- Decision Logging and Audit Trail Tools
- Explainable AI Frameworks (XAI)
- KPMG Trusted AI Framework
- KPMG Trusted AI Risk and Control Matrix (RCM)

Further Reading:

- Google PAIR Guidebook
- OECD AI Principles – Transparency
- AI Explainability 360 (IBM)

Principle 6

Human Oversight

The UAE emphasizes the irreplaceable value of human judgment and oversight over AI, ensuring that AI systems are aligned with ethical values and social standards to correct any errors or biases that may arise.

Understanding the Principle in Real-World Terms

Human oversight ensures that AI systems are not solely responsible for decision-making, but that humans remain in control, especially in high-risk situations. While AI can assist in processing large volumes of data and making decisions based on patterns, human judgment is essential in ensuring these decisions are ethically sound, unbiased, and aligned with societal values.

Real-world examples:

- **Healthcare AI:** In medical AI applications, human doctors must oversee diagnoses and treatment recommendations, ensuring suggestions align with patient well-being.
- **Autonomous Vehicles:** In autonomous driving, human oversight is necessary to intervene in emergency situations and ensure safety protocols are followed.

Embedding This Principle into AI Governance

To ensure effective human oversight, organizations should incorporate mechanisms within their AI governance that enable human intervention and judgment. This includes building systems that allow human supervisors to easily monitor, override, or adjust AI decisions, particularly in situations where decisions could have significant social or ethical implications. Adopting these practices aligns with both the UAE's human oversight principle and international standards, such as the EU AI Act, which emphasizes the importance of human intervention, particularly in high-risk AI applications. Ensuring human oversight can help organizations meet regulatory requirements and avoid the risks associated with fully autonomous decision-making.

Best Practices and Methodologies

Strategy and Design

- **Design for Human-AI Collaboration:** Develop AI systems that prioritize human decision-making, allowing AI to assist but not replace human judgment, especially in high-stakes environments.
- **Create Clear Roles for Human Oversight:** Clearly define the role of human oversight in AI systems, ensuring there are defined protocols for intervention when required.

Data Enablement

- **Bias Detection and Correction:** Implement human-in-the-loop mechanisms to regularly assess and correct biases in training data, ensuring AI models remain fair.
- **Ethical Review of Data:** Conduct ethical reviews to ensure that the data used is aligned with societal values and will not lead to unfair or biased decisions.

Model Development

- **Transparency in AI Decision-Making:** Ensure that the AI model's decision-making process is transparent, enabling humans to understand how decisions are made and intervene if necessary.
- **Human-in-the-Loop (HITL) Systems:** Build systems where human judgment is a key component, especially in high-risk sectors like healthcare, finance, and law enforcement, where human expertise is essential to ensure ethical and fair outcomes.

Testing and Evaluation

- **Human-Centric Testing:** Involve human testers during the evaluation phase to simulate real-world interventions.
- **Audit for Bias and Fairness:** Regularly conduct audits of AI systems to detect any bias, errors, or ethical concerns that may require human correction or intervention.

Deployment and Monitoring

- **Continuous Human Monitoring:** Ensure that human supervisors can monitor AI systems in real-time, especially in critical applications.
- **Emergency Intervention Protocols:** Establish clear protocols for human intervention in case of AI system failures, inaccuracies, or ethical dilemmas, ensuring that human oversight can be swiftly implemented.

Extending the Principle to Agentic AI Systems

Agentic AI challenges traditional oversight due to autonomy and emergent behavior. Organizations must define clear thresholds for human intervention. Design protocols for human-in-the-loop and override scenarios. Oversight mechanisms must include real-time alerts, rollback options, and escalation workflows. Governance teams should conduct periodic reviews of agent performance and decisions. Human accountability should remain visible at all times.

Key Tools, Techniques and Further Reading

Tools and Techniques:

- Human-in-the-Loop (HITL) Platforms
- Amazon SageMaker Ground Truth
- Model Escalation Protocol Templates
- Audit Trail Platforms (e.g. Fiddler AI)
- KPMG Trusted AI Framework
- KPMG Trusted AI Risk and Control Matrix (RCM)

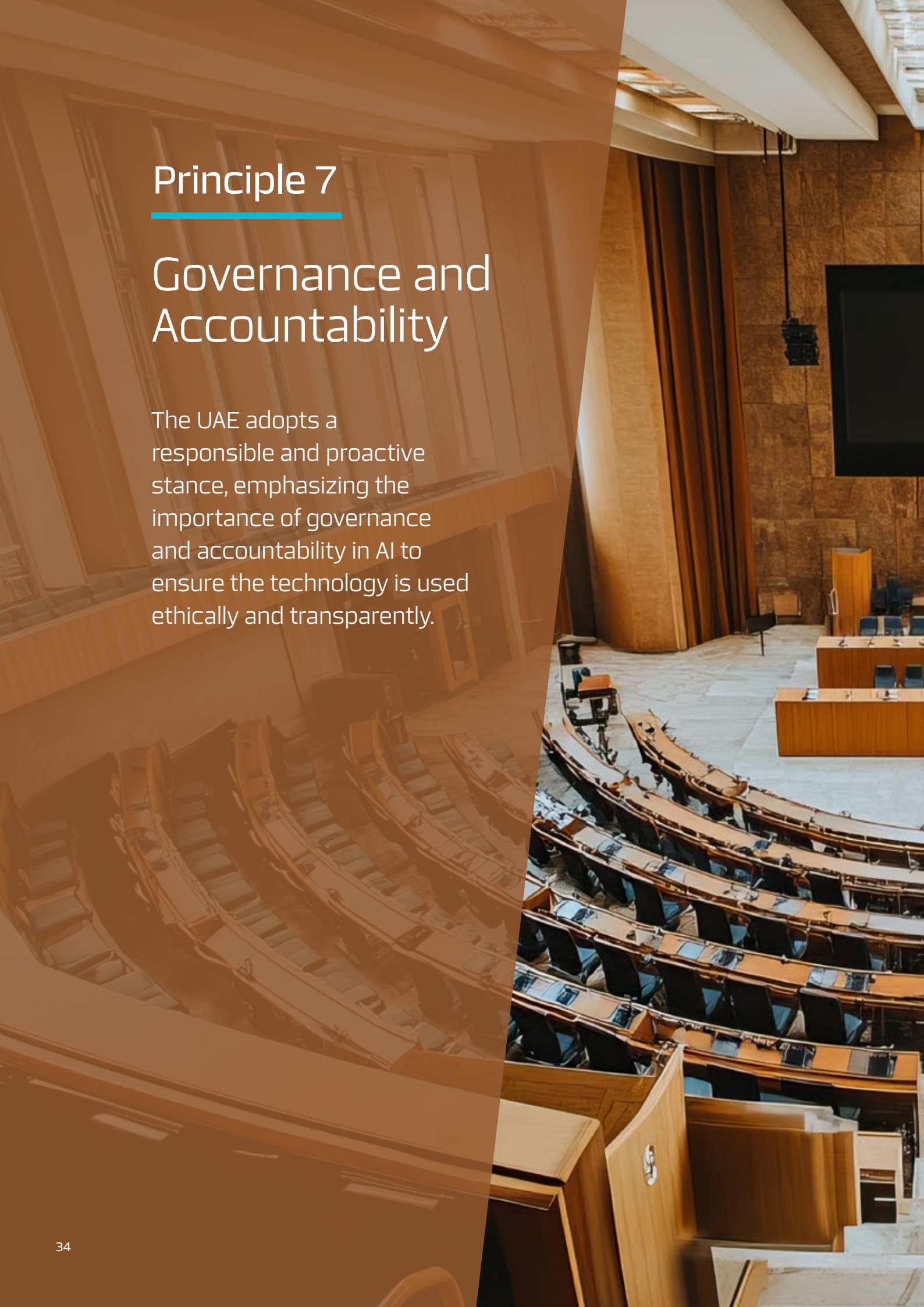
Further Reading:

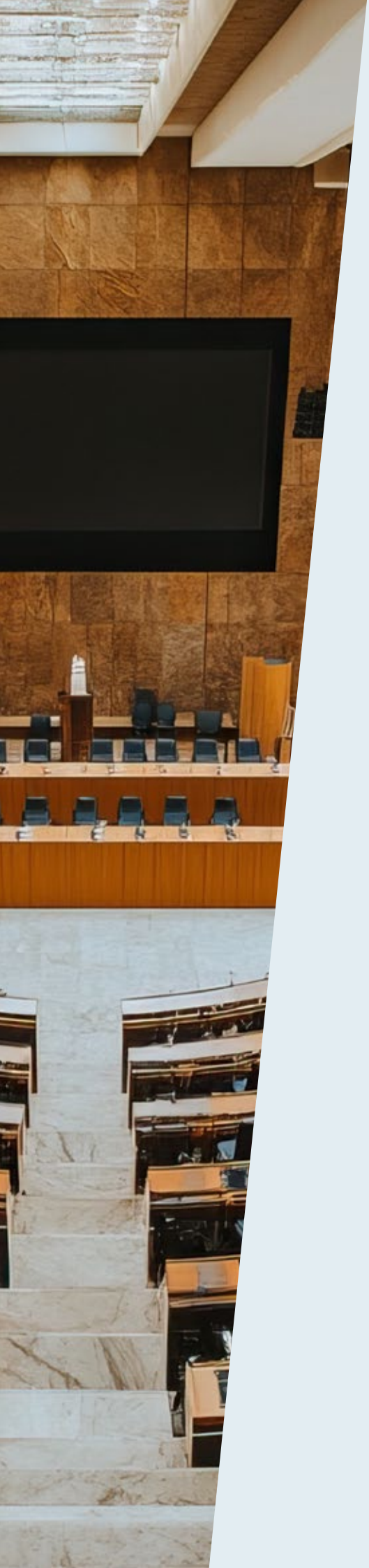
- EU AI Act on Human Oversight
- OECD Principles: Human-Centric AI
- ISO/IEC 42001 AI Management System Draft Standard

Principle 7

Governance and Accountability

The UAE adopts a responsible and proactive stance, emphasizing the importance of governance and accountability in AI to ensure the technology is used ethically and transparently.





Understanding the Principle in Real-World Terms

Governance and accountability in AI refer to establishing clear oversight, ensuring responsible decision-making, and maintaining transparency throughout the AI lifecycle. This involves defining roles and responsibilities, setting clear policies, and implementing mechanisms for monitoring, auditing, and reporting. Without strong governance, AI systems risk becoming opaque or misused, potentially leading to unintended consequences such as discrimination, unfair practices, or breaches of public trust.

Real-world examples:

- **Financial Institutions:** AI systems making loan decisions may lead to biased outcomes if governance structures aren't in place to ensure fairness, transparency, and accountability in the decision-making process.
- **Healthcare:** Without accountability, AI-driven diagnoses could be applied without proper oversight, leading to harmful decisions, damaging patient trust and regulatory compliance.

Embedding This Principle into AI Governance

To effectively integrate governance and accountability into AI systems, organizations in the UAE must develop comprehensive frameworks that define roles, responsibilities, and decision-making processes throughout the entire AI lifecycle. Drawing on global standards such as the EU AI Act, which highlights the importance of supervisory bodies and oversight mechanisms, organizations can establish similar governance structures to ensure accountability and promote the ethical use of AI technologies. By adopting these principles along with best practices, organizations can foster transparency, build trust, and maintain regulatory compliance, positioning themselves as responsible and forward-thinking leaders in AI development within the UAE.

Best Practices and Methodologies

Strategy and Design

- **Define Governance Structures:** Establish clear accountability frameworks that assign roles and responsibilities for AI decision-making and oversight. This should include both technical and ethical oversight committees.
- **Set Ethical Guidelines:** Implement clear ethical guidelines that govern AI use and ensure that these guidelines are followed throughout the AI lifecycle.
- **Design for Human Oversight (HITL):** Assign algorithm owners, define automation levels, correct behavior through input, log interactions externally, and apply MITL for output checks.

Data Enablement

- **Maintain Data Documentation:** Maintain transparent documentation of data sources and decision-making processes to support accountability, making it easy to trace AI outcomes back to the data used.
- **Incorporate HITL:** Embed human oversight in assessing data quality and labeling to enhance reliability and reduce bias.

Model Development

- **Document AI Development Choices:** Maintain clear records of decisions made during model design, including data selection, algorithm choices, and potential ethical considerations.
- **Regular Audits:** Implement ongoing audits to ensure AI systems are operating in line with governance policies and meet transparency standards.
- **Develop HITL Features and Interfaces:** Integrate Human-in-the-Loop mechanisms to ensure oversight, validation, and control in critical decision points.

Testing and Evaluation

- **Accountability Testing:** Test AI systems for accountability by ensuring that the decision-making process can be explained and verified at each step.
- **Reviews:** Conduct regular ethical reviews to assess whether AI systems are being used in a way that aligns with organizational values and societal expectations.

Deployment and Monitoring

- **Continuous Oversight:** Set up mechanisms for ongoing monitoring of AI systems to ensure they remain ethical, transparent, and accountable throughout their lifecycle.
- **Feedback Loops:** Create feedback loops that allow stakeholders, including customers, employees, and regulators, to report any ethical or governance concerns.

Extending the Principle to Agentic AI Systems

Agentic AI requires distributed but auditable governance frameworks. Define clear ownership for agent training, deployment, and adaptation. Implement guardrails and agent behavior policies that align with organizational ethics. Create “agent passports” to track evolution, authority levels, and decisions. Embed agents into existing AI governance boards and review cycles. Establish accountability links between agent decisions and responsible human roles.

Key Tools, Techniques and Further Reading

Tools and Techniques:

- AI Governance Platforms (e.g. IBM Watsonx. Governance, Credo AI)
- Algorithmic Risk Registers
- Model Documentation Templates (e.g. Model Cards)
- Ethical Review Board Guidelines
- KPMG Trusted AI Framework
- KPMG Trusted AI Risk and Control Matrix (RCM)

Further Reading:

- WEF Model AI Governance Framework
- ISO/IEC 38507: Governance of IT – AI Guidance
- UAE AI Ethics and Governance Initiatives

Principle 8

Technological Excellence

AI should be a beacon of innovation, reflecting the UAE's vision of digital, technological, and scientific excellence. The UAE seeks global leadership by adopting technological excellence in AI to drive innovation, enhance competitiveness, and improve quality of life through innovative and effective solutions to complex challenges, contributing to sustainable progress benefiting society.

Understanding the Principle in Real-World Terms

Technological excellence in AI is about continually advancing the capabilities of AI systems to solve pressing challenges, improve efficiency, and drive economic growth. Achieving excellence requires sustained investment in research and development, supported by robust risk and control mechanisms such as KPMG's Trusted AI Framework, which ensures that AI solutions are high-performing while also being ethical, reliable, and secure. The EU AI Act serves as an example by emphasizing rigorous testing, continuous assessment, and clear performance standards to foster innovation while minimizing risks. By embedding trusted frameworks into AI development and deployment, organizations can drive continuous technological improvement while safeguarding the public interest and ensuring regulatory alignment.

Real-world examples:

- **Continuous Model Improvement:** A company regularly updates its AI models with new data and techniques to improve accuracy and performance, ensuring they remain effective and competitive in changing environments.
- **Investment in R&D:** An organization establishes a dedicated AI research lab to explore emerging technologies like generative AI, edge computing, and federated learning to stay at the forefront of innovation.

Embedding This Principle into AI Governance

To embed technological excellence into AI systems, organizations should prioritize continuous innovation, research, and development—while aligning these efforts with ethical standards, regulatory frameworks, and their defined risk appetite. An environment that encourages experimentation and rapid iteration drives breakthroughs but must be underpinned by robust AI risk assessments and a clear risk assessment methodology. These mechanisms enable organizations to innovate confidently while proactively identifying, managing, and mitigating potential risks through well-defined guardrails.

Best Practices and Methodologies

Strategy and Design

- **Foster Continuous Innovation:** Encourage research and development within the organization to explore new frontiers in AI technology.
- **Define Performance Standards:** Establish clear metrics and benchmarks for technological performance, ensuring AI systems meet global standards for excellence.

Data Enablement

- **Ensure High-Quality Datasets:** Use well-curated datasets that support robust and adaptable model development.

Model Development

- **Leverage Advanced Technologies:** Adopt the latest AI algorithms and methodologies to ensure that AI systems remain at the forefront of innovation.
- **Design for Self-Improvement:** Build models capable of continuous learning and adaptation, to maintain technological relevance and performance.

Testing and Evaluation

- **Performance Assessments:** Implement rigorous testing protocols to ensure AI models deliver consistent, high-quality results.
- **Risk and Benefit Evaluations:** Regularly assess the potential risks and benefits of AI deployments to ensure they provide maximum value without unintended consequences.
- **Conduct Comprehensive Testing:** Regularly evaluate AI systems for performance and efficacy to ensure they meet technological excellence standards.

Deployment and Monitoring

- **Ongoing Improvements:** Continue to evolve AI systems post-deployment, incorporating feedback and advances in technology to maintain high standards of performance.

Extending the Principle to Agentic AI Systems

Agentic AI must reflect leading standards in model architecture, interoperability, and resilience. Continuous model tuning, scenario testing, and performance benchmarking are vital. Use advanced simulation environments to model complex agent behavior. Adopt composable architecture to allow modular upgrades. Excellence also includes robustness against adversarial manipulation and unintended consequences.

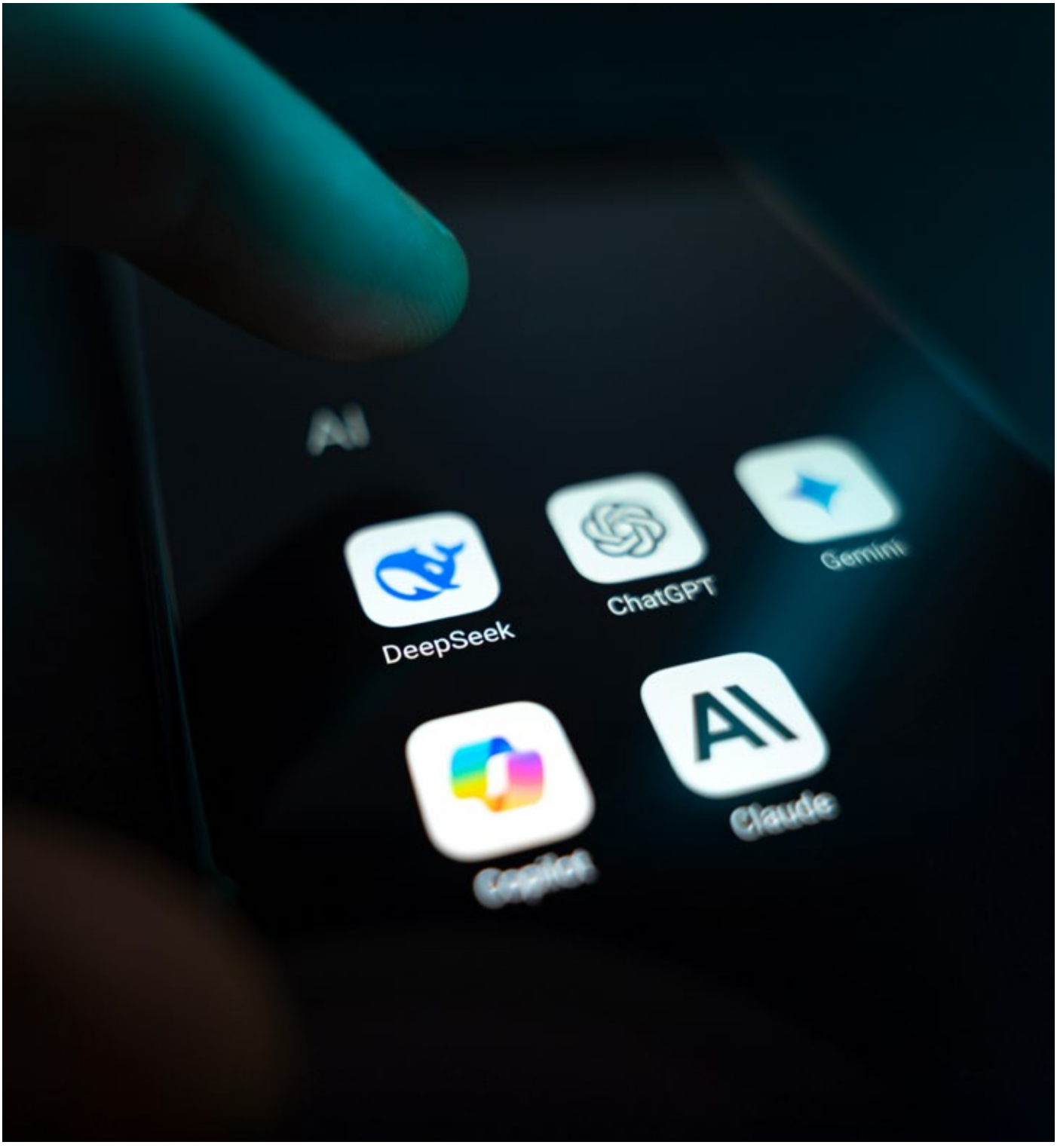
Key Tools, Techniques and Further Reading

Tools and Techniques:

- MLPerf Benchmarking Tools
- CI/CD for ML Pipelines (e.g. MLflow, TFX)
- NVIDIA Triton Inference Server
- Federated Learning Platforms for Edge AI
- KPMG Trusted AI Framework
- KPMG Trusted AI Risk and Control Matrix

Further Reading:

- Stanford AI Index Report
- UAE National AI Strategy 2031
- Microsoft Responsible AI Standard



Principle 9

Human Commitment

Human commitment in AI reflects the spirit of the UAE, essential for ensuring that the development of this technology serves the public good. It focuses on enhancing human well-being and protecting fundamental rights, emphasizing the importance of placing human values at the heart of technological innovation to ensure a positive and lasting impact on society.

Understanding the Principle in Real-World Terms

Human commitment in AI stresses the need to design and deploy AI systems that prioritize human welfare and rights, aligning with the core values and cultural principles of the UAE. This principle advocates for the creation of AI technologies that serve societal interests while safeguarding individual freedom and dignity. The UAE's focus on human commitment ensures that AI innovations benefit all people, creating solutions that promote social good and protect fundamental freedoms.

Real-world examples:

- **AI in Healthcare:** AI-driven solutions that help doctors make better decisions, ensuring better patient care while safeguarding patients' rights and well-being.
- **AI for Disability Access:** AI-powered tools designed to aid individuals with disabilities, ensuring equal access to education, employment, and daily services.

Embedding This Principle into AI Governance

To integrate human commitment into your AI development process, prioritize the well-being of users and society. Ensure that AI systems respect fundamental rights, including privacy, equality, and autonomy. Develop strong governance structures that ensure AI models are inclusive and transparent, focusing on providing tangible benefits to humanity while protecting individuals from potential harm.

Best Practices and Methodologies

Strategy and Design

- **Human-Centered Design:** Develop AI systems that prioritize human needs and well-being throughout the design and development phases.
- **Ethical AI Governance:** Establish clear ethical guidelines that place human rights and societal well-being at the core of AI projects.
- **Public Good Framework:** Align AI development with societal well-being, ensuring that AI solutions contribute positively to the public interest alongside organizational and commercial goals.

Data Enablement

- **Respect for Privacy:** Ensure that data used to train AI models is collected, stored, and processed with full respect for individuals' privacy and rights.
- **Inclusive Data:** Use data that represents diverse demographic groups, ensuring that AI models cater to the needs of all communities, especially marginalized groups.
- **Data Transparency:** Provide clear documentation of data sources and the processes used to ensure the ethical handling of personal and sensitive information.

Model Development

- **Bias Mitigation:** Implement strategies to reduce bias and ensure that AI models do not disproportionately harm or benefit certain individuals or groups.
- **Human Oversight:** Maintain mechanisms for human oversight in the decision-making process, ensuring that AI systems augment human judgment rather than replace it.

Testing and Evaluation

- **Ethical Impact Assessment:** Regularly evaluate the ethical implications of AI systems, including potential harms, bias, and unintended consequences.
- **Human Welfare Testing:** Test AI systems to assess their impact on human well-being, ensuring that they contribute positively to societal progress.
- **Stakeholder Engagement:** Continuously engage with diverse stakeholders, including the general public, to understand the societal impact of AI systems and ensure they align with public interests as well.

Deployment and Monitoring

- **Ongoing Human Rights Monitoring:** Track the deployment of AI systems to ensure they continue to respect human rights and do not inadvertently harm communities.
- **Transparency in Outcomes:** Regularly disclose the outcomes and impacts of AI systems, maintaining public trust and ensuring accountability.
- **Feedback Loops for Improvement:** Implement systems for continuous feedback to ensure that AI systems evolve in ways that enhance human welfare and societal good.

Extending the Principle to Agentic AI Systems

Agents must be designed with human welfare as their ultimate objective. Their operations should uplift accessibility, equality, and well-being across society. Agents must adapt to user preferences, abilities, and cultural values. Ethical constraints should be hardcoded to prevent exploitation or harm. Monitoring systems should assess human impact and surface any ethical concerns early.

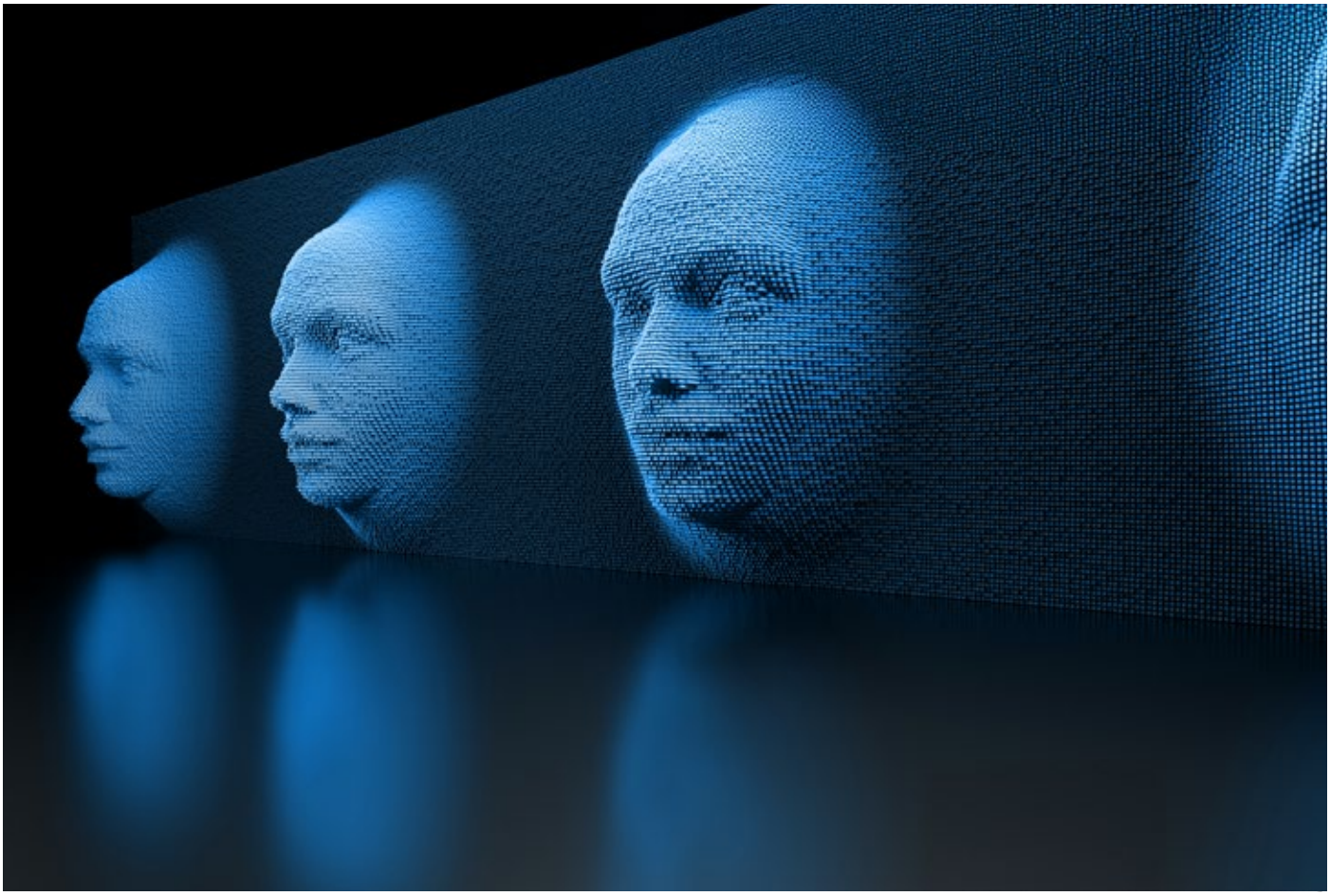
Key Tools, Techniques and Further Reading

Tools and Techniques:

- Human Rights Impact Assessment Frameworks
- DEI-Aware Data Labeling Tools
- Ethical AI Checklists
- Stakeholder Co-Creation Platforms
- KPMG Trusted AI Framework
- KPMG Trusted AI Risk and Control Matrix (RCM)

Further Reading:

- IEEE Ethically Aligned Design
- UN B-Tech Project on Human Rights and AI
- Berkman Klein Center – AI and Civil Liberties



Principle 10

Peaceful Coexistence with AI

The UAE emphasizes peaceful coexistence with AI to ensure that this technology is leveraged to enhance the well-being and progress of communities, without compromising human security or fundamental rights.

Understanding the Principle in Real-World Terms

Peaceful coexistence with AI calls for ensuring that AI systems complement human life and society, enhancing productivity and well-being while maintaining control over technology. This principle focuses on the ethical integration of AI, ensuring that AI systems do not pose threats to human rights, privacy, or security.

Real-world examples:

- **Workplace Automation:** AI systems supporting workers by ensuring the welfare of employees through reskilling and job creation.
- **AI in Surveillance:** AI in surveillance being used in a manner consistent with human rights and security.

Embedding This Principle into AI Governance

To ensure peaceful coexistence, businesses must develop AI technologies that are not only efficient but also ethically aligned with societal values – respecting privacy, human rights, and promoting social cohesion. Ethical deployment should include change management, ongoing training, and KPIs to monitor impact. A framework like KPMG's Trusted AI can reinforce such priorities by embedding ethical considerations, responsible data use, and adaptability into AI systems – helping organizations build trust and stay ahead of evolving regulations, such as the EU AI Act.

Best Practices and Methodologies

Strategy and Design

- **Ethical AI Design:** Prioritize the human impact of AI systems from the start by designing AI solutions that promote fairness, equity, and respect for privacy.
- **Stakeholder Inclusivity:** Involve diverse stakeholders in the AI design process, including ethicists, legal experts, and the communities affected, to ensure alignment with broader societal goals.

Data Enablement

- **Bias Mitigation:** Ensure AI systems are trained on diverse, representative datasets to prevent biased decision-making that could harm vulnerable communities.
- **Privacy Preservation:** Use anonymization and encryption techniques to protect user data, ensuring AI systems do not violate individual privacy or expose sensitive information.

Model Development

- **Build Explainability into Model Architecture:** Prioritize models that naturally lend themselves to interpretability, making it easier for stakeholders to understand, question, and trust AI decisions.

Testing and Evaluation

- **Impact Assessment:** Conduct social and ethical impact assessments before deploying AI systems, evaluating potential risks to human rights, security, and societal harmony.
- **Fairness Audits:** Regularly audit AI systems for fairness, ensuring they do not inadvertently discriminate or violate basic human rights.

Deployment and Monitoring

- **Ongoing Human Oversight:** Implement continuous human oversight of AI systems, especially in high-stakes areas such as law enforcement, healthcare, and financial services, to ensure that AI works in harmony with human values and societal norms.
- **Ethical AI Reporting:** Publish regular reports on the ethical use of AI, detailing any issues or concerns, along with measures taken to address them.

Extending the Principle to Agentic AI Systems

Agentic AI should foster social harmony, not disruption. Ensure that agents respect user boundaries and operate with contextual appropriateness. Avoid designs that encourage dependency or manipulation. Support multi-stakeholder engagement in evaluating agent use in public spaces. Incorporate ethics-by-design to align agents with peaceful coexistence. Monitor agent influence on human behavior and social norms.

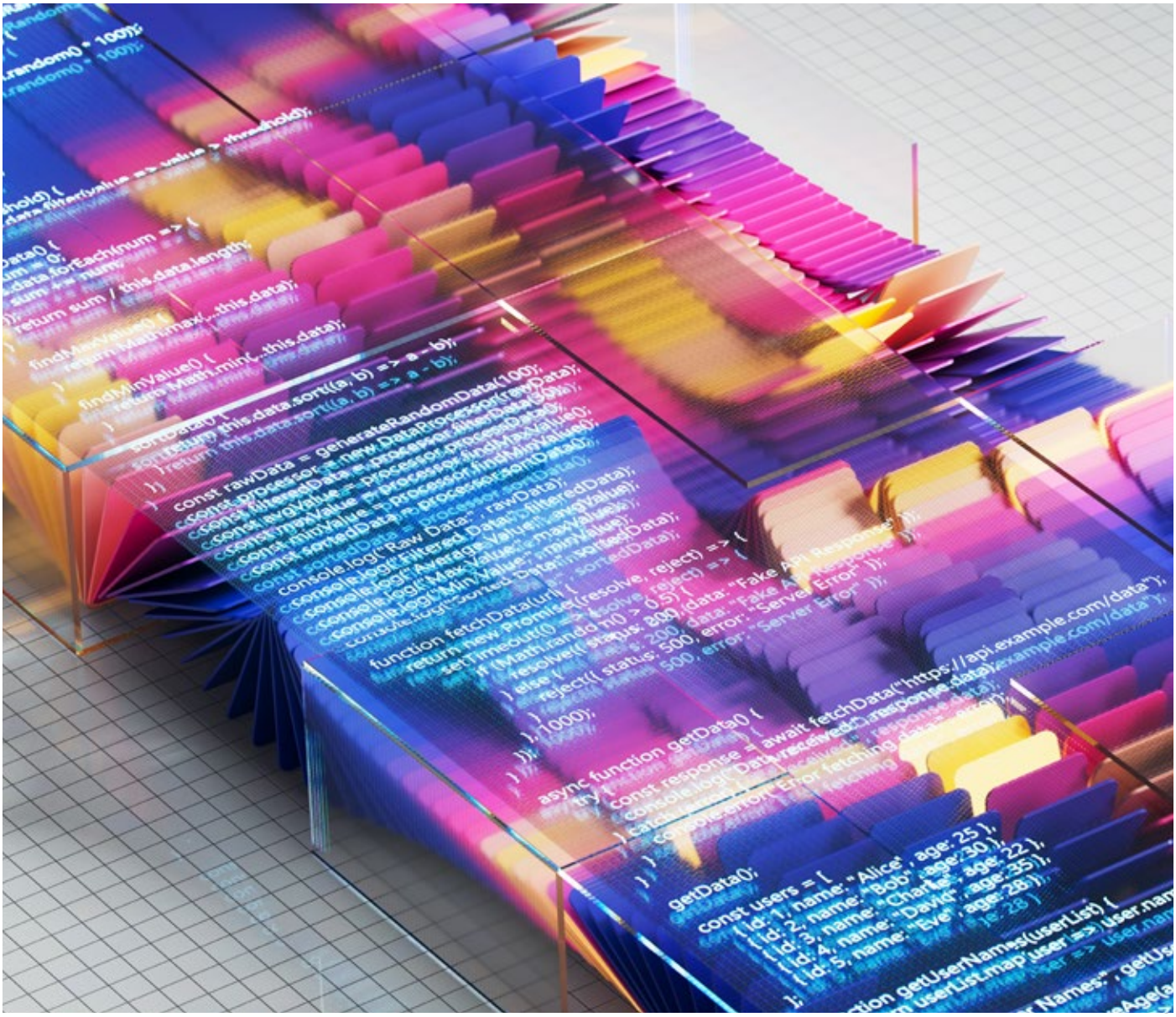
Key Tools, Techniques and Further Reading

Tools and Techniques:

- Ethical Impact Assessment Templates
- Bias Auditing Tools
- Community Feedback Loops (via apps/web interfaces)
- AI for Social Good Design Kits
- KPMG Trusted AI Framework
- KPMG Trusted AI Risk and Control Matrix (RCM)

Further Reading:

- UNESCO AI Ethics Recommendations
- Future of Life Institute – AI Policy Resources
- UAE Ministerial Briefs on AI in Society



Principle 11

Promoting AI Awareness for an Inclusive Future

It is essential to create an inclusive future that ensures everyone can benefit from advancements in AI, guaranteeing equitable access to this technology and its advantages for all segments of society.

Understanding the Principle in Real-World Terms

Promoting AI awareness ensures that everyone—business leaders, students, policymakers, and the general public—can engage with and understand AI technologies. Without widespread literacy and awareness, AI risks becoming an exclusive tool, limiting its benefits to a small group. By fostering AI knowledge and involvement across various demographics, the UAE ensures that AI technologies contribute to societal progress and are inclusive of all people.

Real-world examples:

- **AI Education in Schools:** Introducing basic AI concepts into school curriculums to help students from all backgrounds understand and engage with AI from an early age.
- **Community AI Workshops:** Hosting free public sessions to raise awareness about how AI impacts daily life, targeting groups with limited access to technology or digital literacy.

Embedding This Principle into AI Governance

To integrate this principle, focus on building AI literacy programs, ensuring that AI solutions are inclusive and accessible to all societal groups. Encourage initiatives that foster AI engagement and awareness and ensure that your AI systems are designed to serve all users, particularly those who may have limited access to technology or knowledge. Embrace transparency and provide opportunities for everyone to learn about and benefit from AI advancements.

Best Practices and Methodologies

Strategy and Design

- **AI Literacy Initiatives:** Design AI training programs for businesses, schools, and the general public.
- **Inclusive Outreach:** Engage underserved communities to make AI accessible to a broader audience.
- **User-Centered Design:** Design AI systems that consider varying user needs and provide different levels of interaction based on expertise.
- **Policy Development:** Ensure AI policies promote equitable access and foster an inclusive AI ecosystem.

Model Development

- **Accessibility Features:** Integrate features that make AI solutions more accessible to people with disabilities and non-experts.

Testing and Evaluation

- **Inclusive Testing:** Test AI systems with a diverse user base to identify challenges and ensure fairness.
- **Continuous Engagement:** Use feedback loops from users to refine AI systems, ensuring they remain relevant to all community segments.

Deployment and Monitoring

- **Public Engagement:** Maintain public-facing channels (workshops, webinars) to continuously engage with communities and raise AI awareness.
- **Impact Tracking:** Track how AI solutions are adopted across different demographic groups to ensure that no one is left behind.
- **Fairness Audits:** Conduct audits to assess whether AI systems promote equitable outcomes across different demographic groups.
- **Continuous Improvement:** Regularly update AI solutions based on societal feedback to keep them inclusive and accessible.

Extending the Principle to Agentic AI Systems

Empower users with the knowledge to understand and control agentic AI systems. Build agent interfaces that are accessible, multilingual, and inclusive. Provide training, onboarding, and help features within agent tools. Promote digital literacy programs that demystify agent capabilities. Agents themselves can be designed to promote awareness and inclusivity through explainable interactions.

Key Tools, Techniques and Further Reading

Tools and Techniques:

- AI Literacy Platforms (e.g. Elements of AI, AI4ALL)
- Multilingual NLP Models (e.g. mBERT, BLOOMZ)
- Accessibility APIs (e.g. WCAG Compliance Tools)
- Responsible Design Templates for Low-Tech Users
- KPMG Trusted AI Framework
- KPMG Trusted AI Risk and Control Matrix (RCM)

Further Reading:

- UAE AI Summer Camp Reports
- WEF Future of Jobs Report – AI Upskilling
- OECD Digital Inclusion Toolkit

Principle 12

Commitment to Treaties and Applicable Laws

The UAE emphasizes the importance of complying with international treaties and local laws in the development and use of AI, ensuring that AI technologies are deployed responsibly and legally across borders.

Understanding the Principle in Real-World Terms

The principle emphasizes the necessity for businesses and organizations to ensure that AI technologies are used legally, adhering to both local laws and global treaties. This ensures that AI applications are not only technologically sound but also compliant with regulatory frameworks that safeguard privacy, security, and fundamental rights. Legal use of AI is paramount, and companies must avoid deploying AI technologies in ways that could violate laws or ethical standards.

Real-world examples:

- **Cross-Border Data Transfers:** Ensuring that AI systems comply with international data protection treaties, such as the General Data Protection Regulation (GDPR) in the EU, when transferring personal data across borders.

- **Export Control Regulations:** AI used in sensitive sectors like defense or critical infrastructure that comply with international export control laws to prevent misuse.

Embedding This Principle into AI Governance

To ensure alignment with both local laws and international treaties, businesses should embed legal compliance into the AI development and deployment processes. This includes conducting regular legal reviews and ensuring AI technologies align with the requirements of international treaties, conventions, and other relevant regulatory frameworks, such as the EU AI Act. This act, for example, also underscores the importance of compliance with local laws and international treaties, particularly when it comes to ensuring that AI systems respect fundamental rights and promote public safety. This can help build a clear framework for the alignment of AI systems with global standards. By upholding this principle and following the best practices, organizations will be better positioned to prevent legal issues and remain compliant with evolving AI regulations globally.

Best Practices and Methodologies

Strategy and Design

- **Legal Review in AI Design:** Consult with legal experts before developing AI systems to ensure that systems align with both local and international treaties and applicable laws.
- **Global Compliance Strategy:** Develop a global strategy that ensures compliance with both regional laws (such as data protection laws) and international treaties to avoid legal pitfalls when expanding into new markets.

Data Enablement

- **Cross-Border Data Compliance:** When processing personal data across borders, ensure adherence to international agreements and data protection laws like the GDPR, ensuring that data subjects' rights are respected in all jurisdictions.
- **Data Minimization:** Apply principles such as data minimization to avoid over-collection of data that could violate privacy laws or treaties.

Model Development

- **Legal Audit of AI Models:** Implement regular legal audits of AI models to ensure compliance with relevant laws, including intellectual property rights, privacy protections, and export controls.
- **Human Rights Impact Assessments:** Ensure that AI models do not infringe on basic human rights and comply with both international human rights treaties and national laws protecting such rights.

Testing and Evaluation

- **Legal Risk Assessment:** Regularly evaluate AI systems for legal risks, such as violations of privacy laws, data protection regulations, or the rights of individuals, ensuring compliance with international treaties and local legislation.
- **Legal Compliance Metrics:** Establish measurable metrics to assess the legal compliance of AI systems, and their adherence to local laws.

Deployment and Monitoring

- **Continuous Legal Oversight:** Monitor AI systems post-deployment to ensure they remain compliant with evolving local laws and international treaties. This includes updating AI governance practices to reflect any changes in relevant laws or regulations.
- **Legal Transparency Reports:** Publish transparency reports outlining how AI systems comply with local and international legal frameworks, ensuring accountability.

Extending the Principle to Agentic AI Systems

Agentic AI must comply with evolving laws, including international treaties and jurisdictional regulations. Implement legal compliance modules that track changes in regulatory requirements. Enable agents to adapt their behavior based on local legal context. Maintain logs and audit trails for legal defensibility. Collaborate with legal, compliance, and audit teams to ensure end-to-end oversight. Agents must never act beyond their authorized scope or violate data, safety, or human rights statutes.

Key Tools, Techniques and Further Reading

Tools and Techniques:

- Compliance Management Systems (CMS)
- Legal Risk Scanning Tools (e.g. Relativity AI, NLP-based legal clause flaggers)
- Cross-Border Data Transfer Frameworks (e.g. SCC tools, OneTrust)
- Export Control Compliance Monitors
- KPMG Trusted AI Framework
- KPMG Trusted AI Risk and Control Matrix (RCM)

Further Reading:

- OECD AI Principles
- UAE Cyber Law and Data Regulations Portal
- EU AI Act (Chapters on Legal Accountability)

The Call to Action

The foundation of ethical AI

Clear, actionable AI principles are the bedrock of responsible AI. As KPMG observes, these principles build trust, accountability and innovation. The UAE's AI Charter lays out 12 key principles as a comprehensive governance framework, but only by operationalizing them (turning them into processes and controls) can organizations align AI to societal values.

Framework alignment

The Charter's principles closely align with KPMG's Trusted AI framework, providing practical guidance for implementation. By combining the Charter with Trusted AI, firms gain step-by-step, ethics-by-design guidance that makes AI **systems transparent, explainable** and **compliant.**

From aspiration to action

Companies must move beyond aspirational ethics statements to concrete governance. As one expert notes, "the challenge is moving from the aspirational AI ethics principles… to a space where it's an actuality". In practice, this means building accountability mechanisms and oversight (audit, testing, review boards) that embed the Charter's intent into every AI lifecycle stage.

Benefits of early action

Proactive alignment with the Charter (and pending regulations) pays off. Early adopters build **stakeholder trust** and credibility, **reduce AI risks** (bias, safety, reputational harm) and **spur responsible innovation**, gaining a competitive edge. They also position themselves for **compliance and resilience** – meeting emerging laws (e.g. the EU AI Act) with less disruption.

Actionable Recommendations

Conduct a principles gap and risk inventory

Review current AI projects against the UAE Charter (and draft EU AI Act standards). Catalog systems and classify them by risk to identify governance, privacy or fairness gaps.

Establish robust AI oversight

Set up a cross-functional AI ethics or governance board (including legal, technical and business leaders) to enforce human oversight, accountability and regulatory compliance in all AI initiatives.

Embed ethics-by-design

Integrate privacy protections, explainability and bias mitigation throughout the AI lifecycle. Adopt best practices (e.g. KPMG's Trusted AI framework) so that models are transparent, fair and secure by default.

Prepare for regulation

Inventory and classify AI systems now according to EU risk tiers and begin building required documentation/testing. Early action (such as designating high-risk use cases and writing conformity checklists) will simplify compliance with the EU AI Act and similar laws.

Educate and engage stakeholders:

Train your teams on ethical AI standards and new regulations and communicate your AI policies openly with customers and partners. Clear, ongoing dialogue about AI controls and use-cases builds trust and drives inclusive, responsible adoption.

By internalizing the UAE Charter's principles through a structured governance framework, organizations can turn ethics into action – reducing risk today and laying the groundwork for tomorrow's AI opportunities and requirements.

Next Steps

Organizations that begin this journey of alignment early will be better positioned to adapt to change, demonstrate accountability, and deploy AI in ways that are both effective and ethical. While the approaches presented in this paper provide a strong foundation, AI governance must be tailored to an organization's unique context, strategic priorities, and risk landscape. Moving from principle to practice requires sustained coordination across leadership, oversight functions, and technical execution. KPMG supports this transformation through its globally recognized Trusted AI Framework, complemented by tools such as the AI Risk and Control Matrix (RCM) and the Governance Operating Model. Together, these offerings enable organizations to assess AI maturity, design tailored governance structures, implement robust risk controls, foster a culture of responsible innovation, and ultimately align with both local and global AI principles and regulatory requirements. For organizations ready to act, KPMG brings cross-disciplinary expertise to embed AI governance that is future-ready.

KPMG Middle East LLP

KPMG Middle East LLP is a part of the KPMG global organization of independent member firms that operate in 143 countries and territories and are affiliated with KPMG International Limited. We provide audit, tax and advisory services to public and private sector clients across Saudi Arabia, United Arab Emirates, Jordan, Lebanon, Oman, and Iraq, contracting through separate legal entities. We have a strong legacy in the region, where we have been established for over 50 years. KPMG Middle East LLP is well-connected with its global member network and combines its local knowledge with international expertise.

KPMG serves the diverse needs of businesses, governments, public-sector agencies, not-for-profit organizations, and the capital markets.

Our commitment to quality and service excellence underpins everything we do. We strive to deliver to the highest standards for our stakeholders, building trust through our actions and behaviors, both professionally and personally.

Our values guide our day-to-day behavior, informing how we act, the decisions we make, and how we work with each other, our clients, and all our stakeholders. Integrity: We do what is right. Excellence: We never stop learning and improving. Courage: We think and act boldly. Together: We respect each other and draw strength from our differences. For Better: We do what matters.

Our purpose is to inspire confidence and empower change. By inspiring confidence in our people, clients and society, we help empower the change needed to solve the toughest challenges and lead the way forward.

KPMG's Our Impact Plan guides our commitments to serving our clients, people and communities across four categories: Planet, People, Prosperity, and Governance. These four priority areas assist us in defining and managing our environmental, social, economic and governance impacts to create a more sustainable future. We aim to deliver growth with purpose. We unite the best of KPMG to help our clients fulfil their purpose and deliver against the United Nations Sustainable Development Goals, so all our communities can thrive and prosper.

We are dedicated to delivering growth with purpose, helping our clients achieve their goals, and advancing sustainable progress to ensure that all our communities thrive. Empowered by our values, and committed to our purpose, our people are our greatest strength. Together, we are building a values-led organization of the future. For better.

Disclaimer: Some or all of the services described herein may not be permissible for KPMG audit clients and their affiliates or related entities.

Contacts

Emilio Pera
Deputy CEO- Middle East, CEO- Lower Gulf
KPMG Middle East
emiliopera@kpmg.com

Robert Ptaszynski
Head of Digital & Innovation and Managed Services
KPMG Middle East
rptaszynski@kpmg.com

Matin Jouzdani
Partner, Data, Analytics and AI
KPMG Lower Gulf
mjouzdani1@kpmg.com

Joe Devassy
Director, Strategic Alliances
KPMG Lower Gulf
jdevassy@kpmg.com

WORLD GOVERNMENTS SUMMIT

JOIN THE CONVERSATION

@WorldGovSummit

www.worldgovernmentssummit.org