Aldar Audit Bureau
Abdullah Al-Basri & Co.
Member firm of Grant Thornton International Ltd.
Grant Thornton

Strengthening Digital Trust in Saudi Arabia's AI-Driven Economy

Insights from the World Economic Forum's 2025 Emerging Technologies Report with a Grant Thornton Perspective.

Contents

Strengthening Digital Trust in Saudi Arabia's AI-Driven Economy

Insights from the World Economic Forum's 2025 Emerging Technologies Report with a Grant Thornton Perspective.

Abdulmajid Al-Basri, CEO
Aldar Audit Bureau, Grant Thornton Saudi Arabia

Artificial intelligence is no longer an emerging concept confined to research laboratories or niche applications. It has become the backbone of global digital transformation, shaping industries as diverse as healthcare, financial services, logistics, and entertainment. Generative AI in particular has leapt into the mainstream, producing vast amounts of synthetic text, images, audio, and video with levels of realism that would have been unthinkable just a few years ago. This transformation offers extraordinary opportunities for productivity, creativity, and innovation. At the same time, it poses one of the most pressing challenges of our digital era: how to maintain trust in information when the boundary between human and machine-generated content is increasingly blurred.

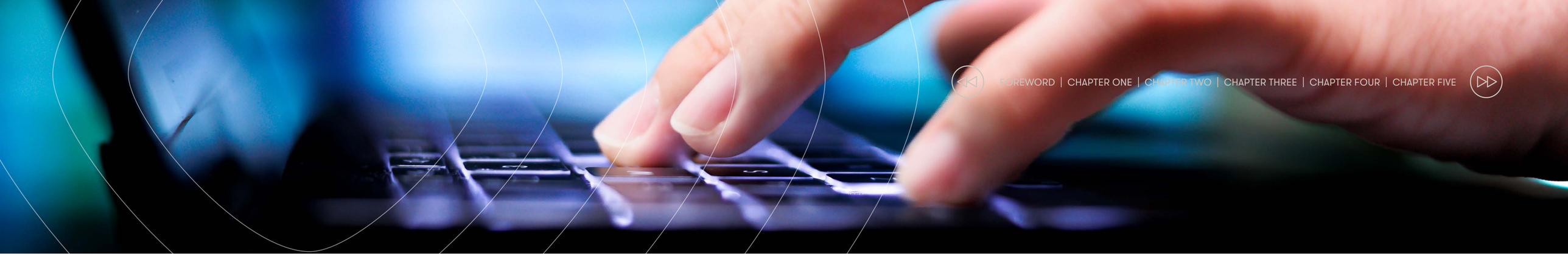
The World Economic Forum's (WEF) Top 10 Emerging Technologies of 2025 places this challenge at the centre of its analysis. Among the selected breakthrough technologies is generative watermarking a system of embedding invisible and tamper-resistant signals into AI-produced content so that its origin can be authenticated and traced. The report emphasises that watermarking is not simply a technical innovation but a governance tool. It is an emerging pillar of what could become a global framework for digital trust. For businesses, governments, and societies, watermarking represents a defence against misinformation and fraud, a safeguard for intellectual property, and ultimately a foundation for confidence in digital economies.

For Saudi Arabia, the relevance could not be greater. As the Kingdom drives forward its Vision 2030 strategy, artificial intelligence has been identified as a catalyst for economic diversification, public sector reform, and the development of entirely new industries. The establishment of the Saudi Data and Artificial Intelligence Authority (SDAIA), the creation of AI-driven giga-projects such as NEOM, and the rapid growth of fintech and e-commerce platforms are all part of a national ambition to place AI at the heart of its transformation. Yet the success of this ambition depends on more than technical capacity or financial investment. It depends on citizens trusting government services, investors trusting financial platforms, and global partners trusting that Saudi Arabia's digital infrastructure is reliable and secure.

This article first examines the WEF's findings on watermarking and digital trust and then explores how these insights can be applied to the Saudi context. It concludes with a Grant Thornton perspective on how the Kingdom might build a distinctive leadership position in this field, balancing global alignment with local priorities and ensuring that its AI-driven economy rests on firm foundations of authenticity and integrity.

Chapter 1:

Generative Watermarking: A WEF Perspective.



The WEF defines generative watermarking as the process of embedding imperceptible, durable, and machine-readable signals into AI-generated outputs. These signals act as digital fingerprints, allowing the origin of content to be traced and authenticated. Unlike visible watermarks that can be cropped or obscured, these next-generation marks are invisible to the naked eye yet resilient against most forms of alteration. They can be applied to text, still images, video, or audio.

The urgency of watermarking arises from the explosive growth of generative AI and the mounting risks it carries. Without reliable authentication systems, societies face a proliferation of synthetic misinformation, erosion of intellectual property rights, and widespread uncertainty about the provenance of information. The WEF report points to the growing role of synthetic content in fraud, disinformation campaigns, and market manipulation. It cites the 2025 Global Risks Report, which highlights misinformation as one of the top global threats. Against this backdrop, watermarking is positioned as a cornerstone of digital trust infrastructure, allowing users, regulators, and courts to distinguish between authentic and fabricated content.

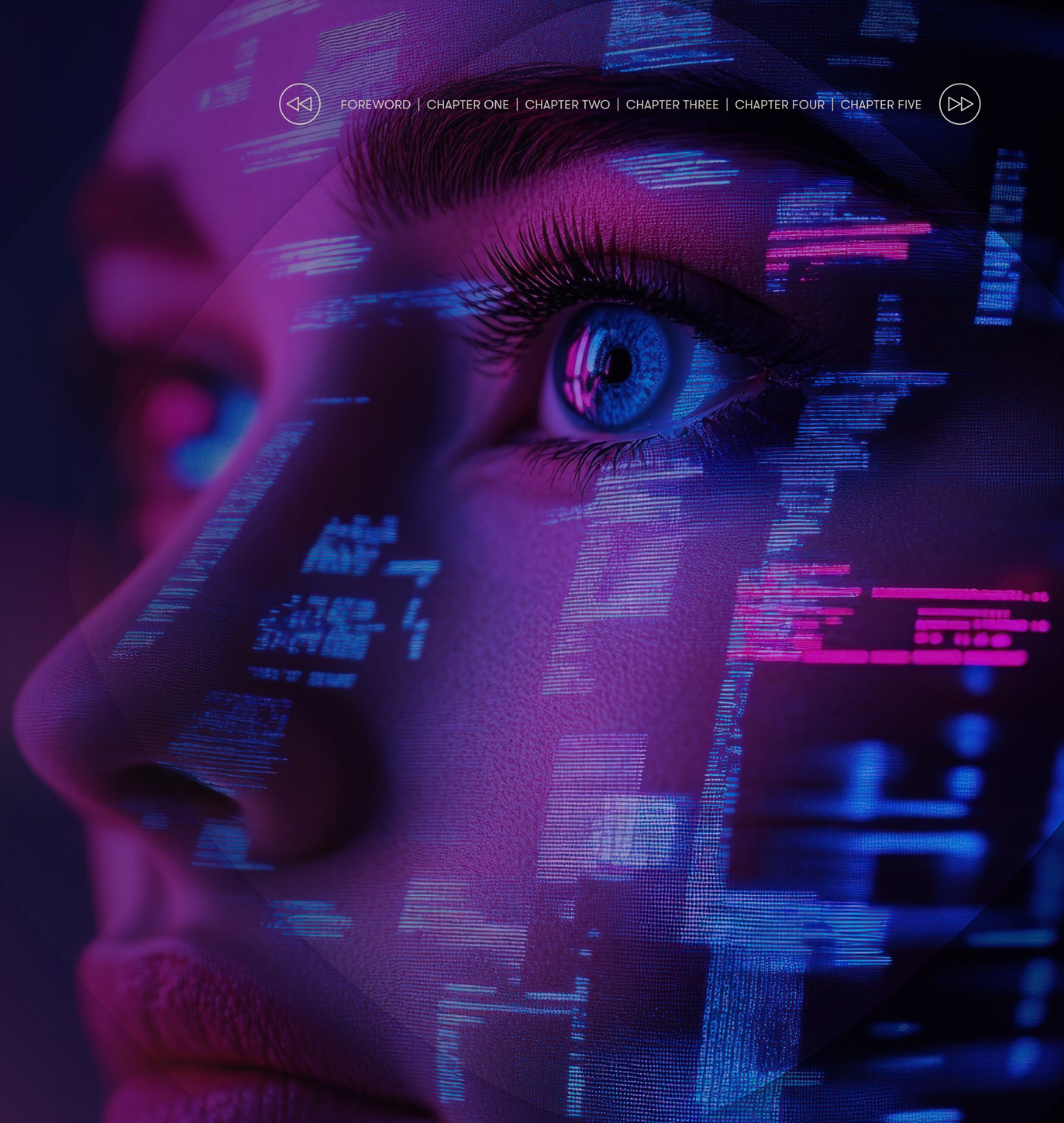
Several initiatives already point to how watermarking may be deployed. Google DeepMind has launched SynthID, which embeds watermarks into images and text produced by its models. Meta has introduced VideoSeal, applying similar techniques to video. Adobe has advanced its Content Credentials initiative, embedding metadata that records the provenance of digital works. Collectively, these efforts demonstrate both the technical feasibility and the commercial interest in watermarking. Yet the WEF cautions that global adoption is uneven, standards are fragmented, and technical robustness is still evolving.

The report also stresses that watermarking is more than a tool for compliance. It is an enabler of new markets and new governance systems. Courts may begin to recognise watermarks as evidence in intellectual property disputes. Insurers could differentiate between content that carries verifiable markers and content that does not, establishing premiums for authenticity. Media companies may find new commercial value in offering “verified content” to audiences seeking trustworthy information. And at the global level, watermarking could intersect with blockchain technologies to provide a distributed system of provenance, authenticity, and accountability across multiple platforms and jurisdictions.

At the same time, the WEF highlights significant challenges. Technical vulnerabilities remain, since watermarks can sometimes be removed or corrupted by adversarial techniques. Lack of interoperability between different systems risks undermining their effectiveness. There is also the risk of false positives, where human-generated content might be incorrectly flagged as synthetic. And adoption is inconsistent: while China already mandates watermarking for AI content, other jurisdictions have yet to legislate. The WEF therefore concludes that watermarking should be viewed as an emerging but incomplete foundation, a critical building block in what will ultimately need to be a comprehensive digital trust architecture.

Chapter 2:

Saudi Arabia's AI-Driven Economy and the Imperative of Trust.



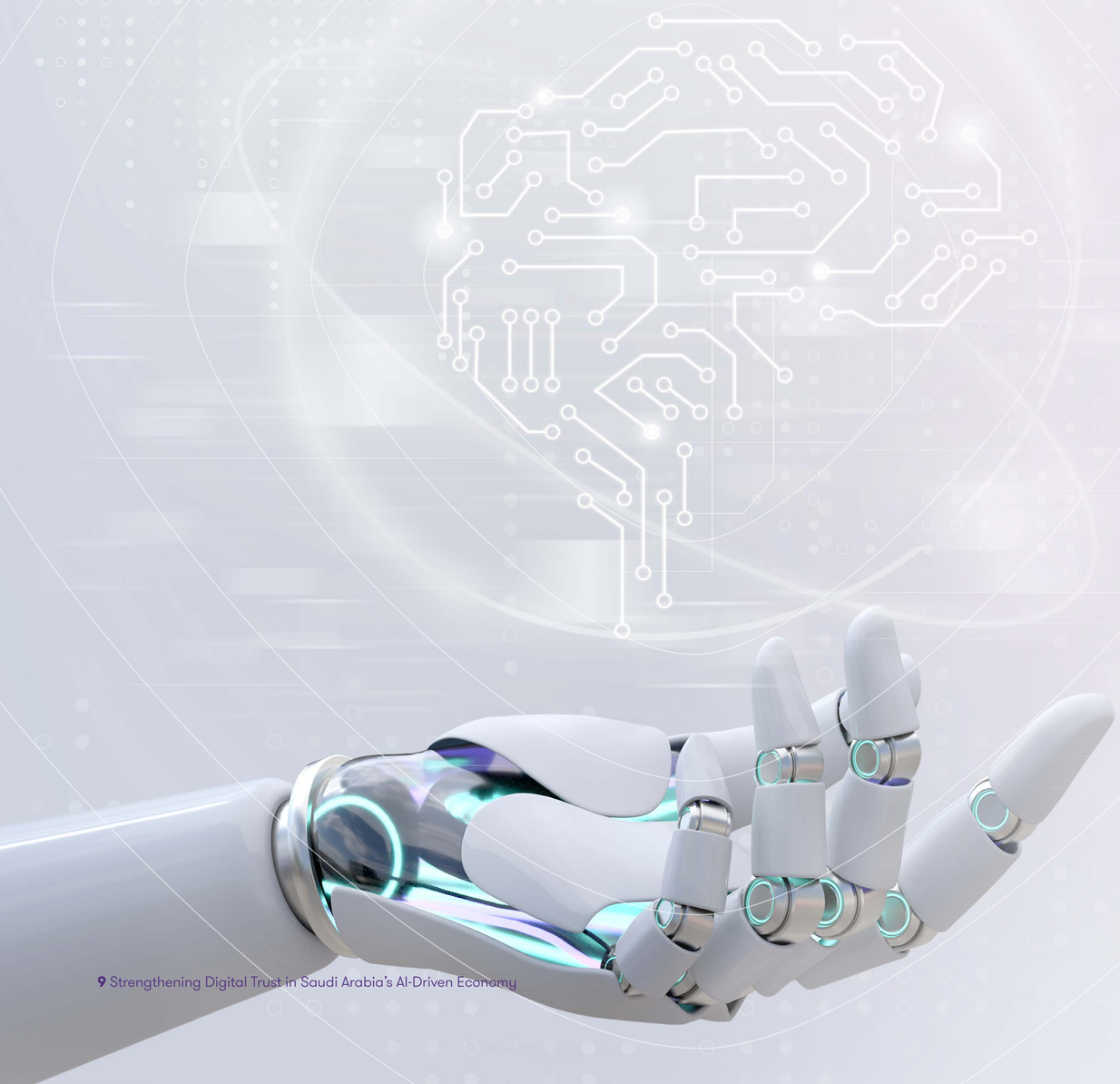
Saudi Arabia is advancing rapidly to establish itself as a regional AI powerhouse. Through Vision 2030, the Kingdom has positioned AI not simply as a supporting technology but as a strategic lever for economic diversification, industrial competitiveness, and public sector efficiency. SDAIA has launched national AI strategies, invested in research partnerships, and created regulatory sandboxes to accelerate adoption. Giga-projects such as NEOM are embedding AI in urban planning, environmental monitoring, and citizen engagement. The financial sector is deploying AI for fraud detection, credit scoring, and regulatory compliance. The cultural and creative industries are exploring AI in gaming, film, and media.

In this context, the question of trust takes on heightened significance. If citizens cannot be confident that government communications are authentic, the legitimacy of public digital services may be undermined. If financial institutions cannot verify the integrity of data or trading signals, the credibility of capital markets may be threatened. If global investors perceive Saudi Arabia's AI environment as unregulated or vulnerable to fraud, capital inflows may slow. Conversely, if the Kingdom can demonstrate robust systems of digital trust, it will enhance its attractiveness as a safe and reliable hub for AI-driven innovation.

Generative watermarking, therefore, aligns closely with Saudi priorities. It offers a mechanism to safeguard markets, protect citizens, and build confidence in the authenticity of digital interactions. It could support financial institutions in preventing fraud, provide courts with reliable evidence in cases involving digital content, and give media and creative industries tools to protect intellectual property. It also resonates with cultural values of honesty, accountability, and transparency, which are integral to the Kingdom's governance traditions.

Chapter 3:

Strategic Application.



The opportunities for watermarking extend across multiple sectors of the Saudi economy.

In the financial sector, the rapid growth of fintech platforms and the ambition to expand Riyadh as a regional financial hub create a pressing need for secure digital verification. Watermarking could play a role in preventing manipulation of synthetic financial reports or AI-generated trading signals. It could enhance compliance with international anti-money laundering and know-your-customer standards by ensuring that digital identities and documents carry authentic verification markers. Courts and regulators could recognise watermarked content as admissible evidence, strengthening the legal framework for digital transactions.

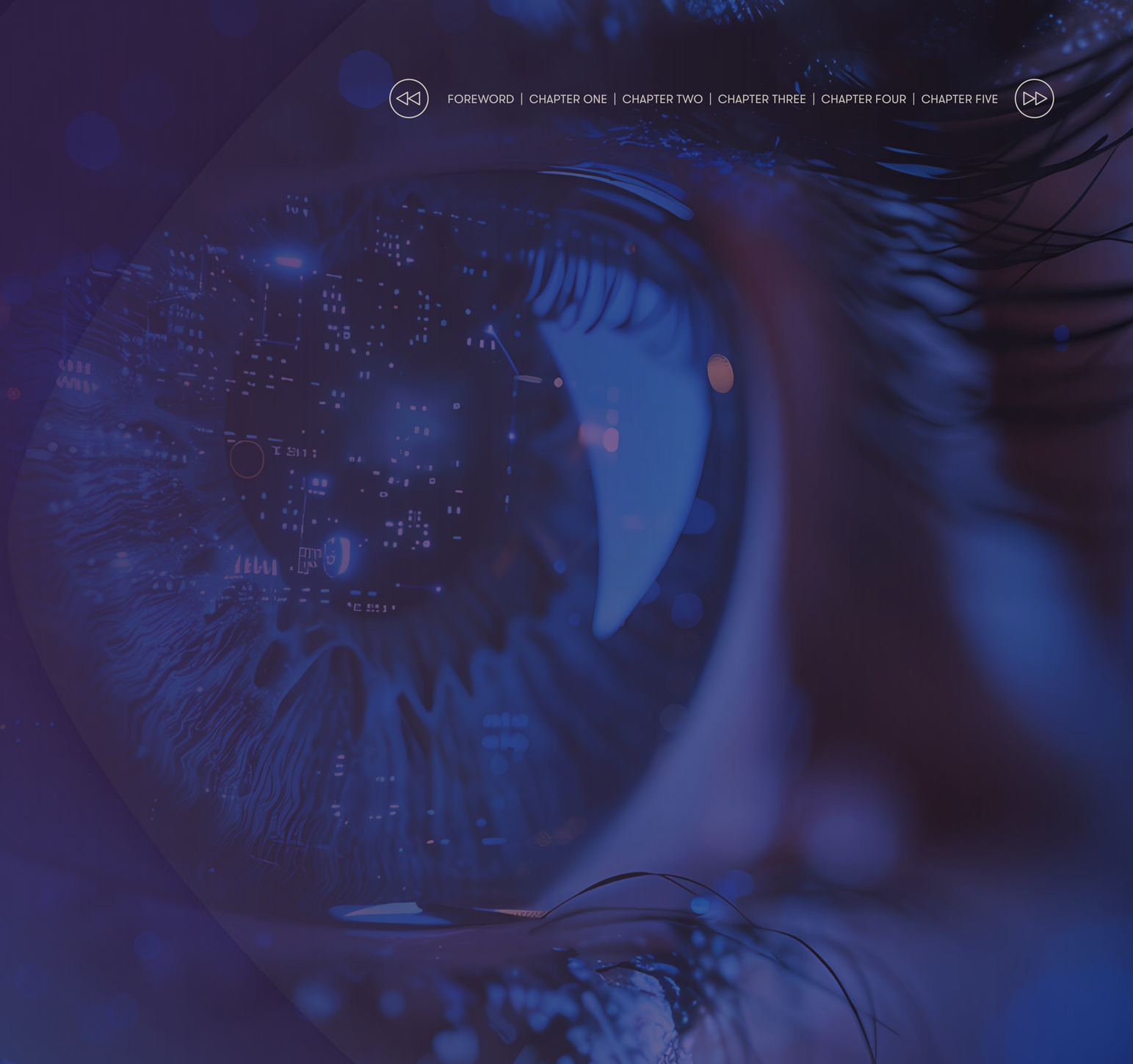
In giga-projects and smart cities, where AI will be embedded in the daily fabric of life, watermarking could ensure that citizen feedback systems are not manipulated by automated bots, that digital twins of infrastructure are not corrupted by tampered data, and that healthcare and education platforms deliver trusted information. In NEOM, which aims to position itself as the world's leading cognitive city, embedding watermarking at the infrastructure level could serve as a differentiator, ensuring that every piece of synthetic data is verifiable and trustworthy.

In the cultural and creative industries, watermarking could provide a mechanism to protect Saudi creators in film, gaming, and digital content from intellectual property theft. It could also enable new commercial models, where verified Saudi content carries a premium in global markets. This aligns with the Kingdom's ambition to expand its entertainment and creative sectors, both as domestic growth engines and as exports.

Even in academic and research institutions, watermarking could protect academic integrity by distinguishing between human and AI-generated work. Universities could adopt watermarking tools to safeguard research, prevent plagiarism, and uphold international standards of scholarship.

Chapter 4:

Grant Thornton Perspective: A Strategic Path Forward.



From Grant Thornton's perspective as an advisor to both the public and private sectors, watermarking and digital trust present both a challenge and an opportunity. The challenge is that trust cannot be retrofitted, it must be built into systems from the outset. The opportunity lies in Saudi Arabia's still-formative AI ecosystem, allowing it to embed digital trust into its strategy's foundations rather than attempting to graft it on later.

The first imperative is regulatory leadership. Saudi Arabia can establish national standards for watermarking and digital trust, aligning them with global frameworks such as the EU AI Act and the C2PA standards, while also tailoring them to local needs. By moving early, the Kingdom can position itself as a regional standard-setter, ensuring that MENA markets converge around its frameworks. This would not only protect domestic interests but also enhance Saudi influence in shaping international governance.

The second imperative is market development. Government incentives could encourage Saudi fintechs, AI startups, and media companies to adopt watermarking technologies. Certification schemes could be created to signal authenticity, similar to the way halal certification signals compliance with religious standards. This would provide consumers and businesses with confidence that content and transactions are genuine, while also creating commercial value for firms that invest in trust.

The third imperative is institutional adoption. Sovereign wealth funds, regulators, and courts can lead by example. The Public Investment Fund could require digital trust mechanisms in its investments. The Capital Market Authority and the Saudi Central Bank could embed watermarking into their compliance systems. Courts could set precedents by recognising watermarks as legal evidence. Universities could use them to uphold academic standards. Such adoption by major institutions would set powerful signals to the market.

The fourth imperative is international collaboration. Saudi Arabia cannot build digital trust in isolation. It will need to partner with global technology firms, align with GCC neighbours, and use its diplomatic influence through platforms such as the G20 and the World Economic Forum to advocate for harmonised standards. By positioning itself as a proactive collaborator, the Kingdom can ensure that its interests are embedded in global frameworks.

Chapter 5:

Looking Ahead: A Saudi Model of Digital Trust.

What emerges from this analysis is the possibility of a distinctive Saudi model of digital trust. This model would be characterised by integrated governance, where regulators, sovereign wealth funds, and ministries work together; by cross-sector application, where financial services, giga-projects, healthcare, and culture all adopt consistent frameworks; by cultural resonance, linking transparency with Islamic principles of accountability and honesty; and by economic advantage, positioning Saudi Arabia as a safe harbour for AI-driven investment and innovation.

Such a model would also reinforce Vision 2030's broader ambitions. By embedding trust into digital transformation, Saudi Arabia can not only accelerate adoption but also ensure sustainability. It can demonstrate to global investors that its markets are not only dynamic but also reliable. It can provide citizens with confidence that AI will enhance rather than undermine their lives. And it can present itself internationally as a voice of integrity in a digital landscape that is often clouded by doubt.

The World Economic Forum's identification of generative watermarking as one of the most important emerging technologies of 2025 highlights a profound reality: without trust, AI economies cannot thrive. For Saudi Arabia, the stakes are particularly high. As the Kingdom seeks to establish itself as a regional AI leader and a global hub for digital innovation, the ability to guarantee authenticity, security, and accountability will determine the credibility of its transformation.

Grant Thornton's perspective is that the Kingdom should not wait for global standards to be imposed. Instead, it should move decisively to set its own benchmarks, align with international frameworks, and embed watermarking into the fabric of its AI strategy. By doing so, Saudi Arabia can transform a technological necessity into a strategic advantage-strengthening its markets, protecting its citizens, and positioning itself as a trusted leader in the global AI economy.

In a world where synthetic media increasingly blurs the line between truth and falsehood, Saudi Arabia has an opportunity to become not only a consumer of trust technologies but a creator of trust standards. By embracing watermarking and digital trust at the highest level, the Kingdom can ensure that its AI-driven economy is built not on shifting sands but on the solid foundation of authenticity and integrity.

Riyadh

U Walk, C4

2nd Floor

Prince Turki Ibn Abdulaziz

Al Awal Road

Riyadh

+966 (11) 463 0680

Jeddah

Crystal Tower,

2nd Floor

King Fahad Rd,

Al Bawadi District,

Jeddah 23443

Saudi Arabia

+966 (12) 691 6883

Dammam

7th Floor

Ababtain Tower

Custodian of the Two

Holy Mosques Road

Al-Khobar

Kingdom of Saudi Arabia

+966 (13) 896 0592

Aldar Audit Bureau

Abdullah Al-Basri & Co.

Member firm of Grant Thornton International Ltd.

Grant Thornton

grantthornton.sa