



How synthetic worlds can help organizations rehearse AI decisions

Why the next act of synthetic data is a place to rehearse decisions, not a better copy of the data you already have.



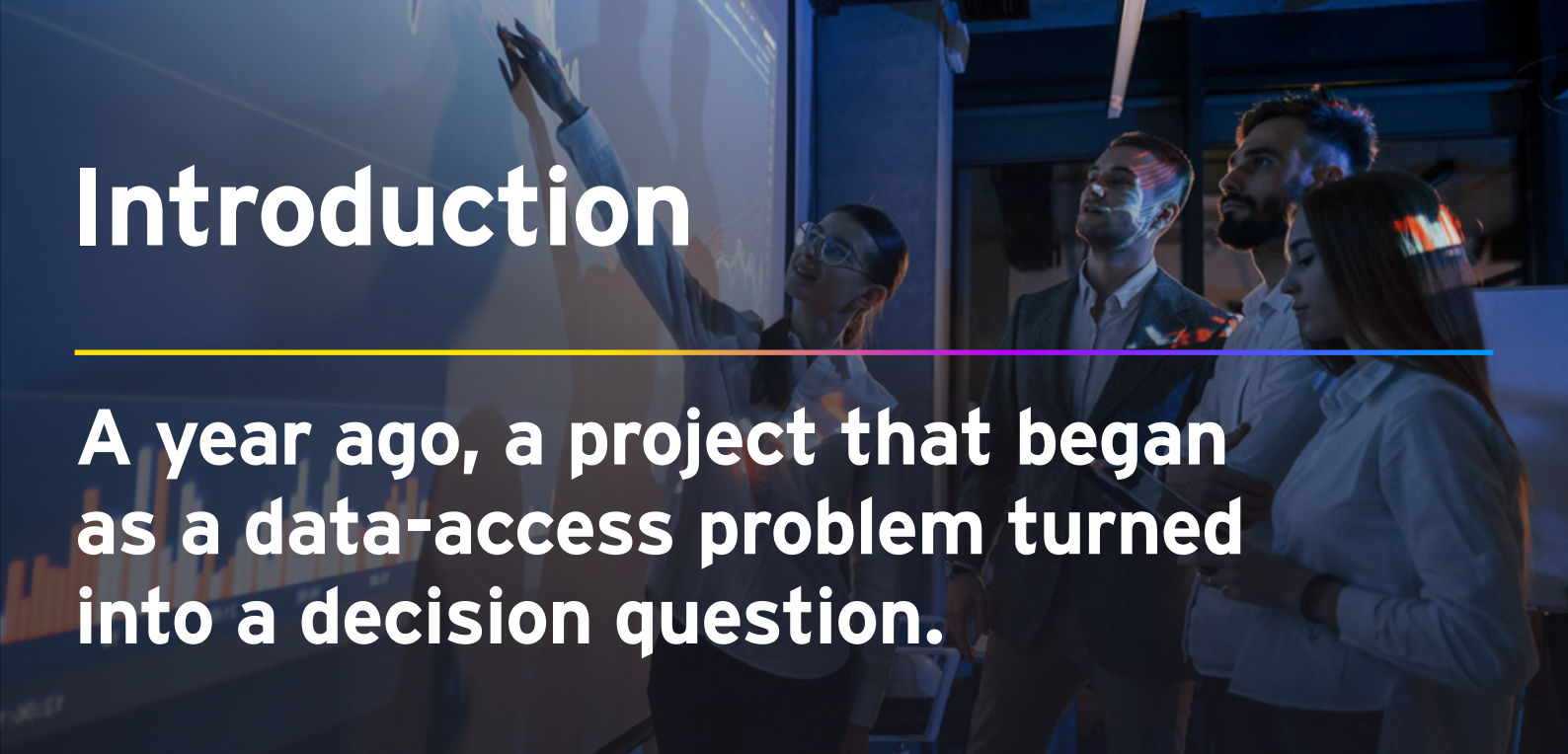
The
Reimagination
Engine



Shape the future
with confidence

Contents

Introduction	3
1 When synthetic data is not enough	5
2 What changes when records become worlds	6
3 A decision rehearsed, not just a dataset copied	6
4 The dangerous side of plausibility	8
5 The evidence frontier: what a synthetic world can and cannot support	10
6 Why AI agents need worlds to rehearse in	12
7 What leaders should do – and avoid	13
8 The disciplined ambition	14



Introduction

A year ago, a project that began as a data-access problem turned into a decision question.

A global humanitarian organisation needed realistic person and household data that teams could use without exposing beneficiary information. Synthetic records helped with that first problem. They gave developers a safer way to build and prototype. But as AI moved closer to live workflows, a harder decision question surfaced: what would happen when an AI-supported decision met changing conditions? If supply was disrupted, eligibility rules changed or an agent routed a case too confidently, where would the process break? What evidence would the organisation need before any real person felt the effect?

The limit showed up quickly. Historic records could support a familiar test, but they froze the problem at the level of records. They did not capture the operating context that can make a decision safe or unsafe once conditions change.

As the work developed, the question shifted from “Can teams build safely against realistic person and household data?” to “Can leaders see the consequences of changing the decision?” The useful signals were not only in the records. They were in the queues that delayed a decision, the exceptions that required a human, the eligibility thresholds that created borderline cases and the evidence labels that showed what could and could not be trusted. That was the moment the work expanded beyond synthetic data into a synthetic world.

“

Leaders need more than a better copy of yesterday's data. They need a world in which they can rehearse tomorrow's decisions.

That is the shift this article is about. The first wave of synthetic data solved a genuine problem: sensitive production data is hard to move or share, so organisations learned to create privacy-safe substitutes for testing, benchmarking and model training. That work still matters. It is simply no longer enough for the decisions leaders now face with AI.

The central thesis is simple: the important shift is not from real data to synthetic data. It is from data production to decision rehearsal. A synthetic world becomes strategically useful only when it preserves the mechanisms relevant to a decision and clearly identifies where simulation ends and real-world evidence must begin.

We call the governed capability for creating and using those environments Synthetic World Intelligence. It moves synthetic data from record-level mimicry to world-level simulation: controlled synthetic realities that can be inspected, challenged and used to test decisions under assumptions that are stated out loud.

For decision-rehearsal use cases, success cannot be measured only by how closely synthetic data resembles reality. Rows are not worlds, and plausibility is not evidence. The more important question is whether an environment can help an organisation reason before it acts – and identify where real-world evidence is still required.



Idea in brief

The problem

The first wave of synthetic data solved privacy-safe substitution: records that resemble real data well enough to test, train or share. Useful, but not enough to tell leaders how a policy, workflow or AI agent will behave before it is switched on.

The insight

Synthetic data becomes strategically valuable when it moves beyond record generation and gives leaders a governed environment in which to rehearse decisions, test assumptions and see where real-world evidence is still needed.

The opportunity

Realistic data generation is becoming table stakes. The defensible capability is governed decision rehearsal: worlds with business logic, stress conditions, known internal truth, and evidence boundaries that senior leaders can challenge.

1

When synthetic data is not enough

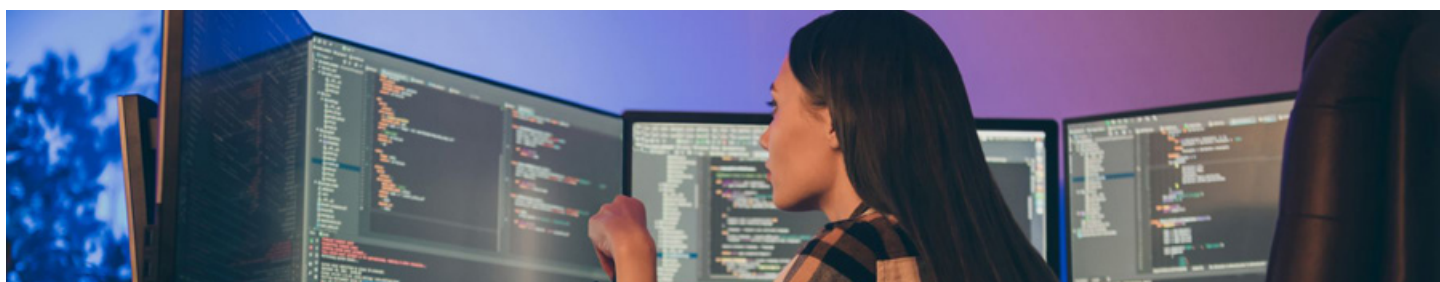
Synthetic data initially gained traction as a way to address data scarcity and privacy constraints. Today, demand is increasingly shaped by a different challenge: institutions need safe environments to test decisions that are too expensive, too risky, too rare or too far in the future to evaluate in the real world.

Several developments are driving this shift. As AI agents move from advice and summarisation toward executing workflows, organisations need greater confidence in how decisions will perform under different conditions. At the same time, increasing volatility is making historical averages a poor guide to future outcomes, while regulation increasingly demands that AI-enabled decisions be explainable and auditable.

Together, they change what a synthetic asset is worth. Its value has traditionally been measured by how closely it resembled an input dataset. For decision-rehearsal use cases, it is also measured by whether it helps leaders reason about mechanisms, stress outcomes and see what real evidence they still need before acting.

Research and industry analysis point in the same direction. The World Economic Forum has argued that synthetic data can fill gaps, protect privacy and test new scenarios, while warning that the line between real and artificial can blur without strong governance and provenance.¹ Gartner has linked the rise of AI agents and automated decisions with synthetic-data management risks around governance, accuracy and compliance.² Together, these developments point away from substitution alone, toward decision infrastructure.

It is important to clarify the scope. Seed synthetic data is still often the right tool when representative source data exists but cannot be used directly. This article is about the next stage. While seed synthetic data starts from the data you happen to have, synthetic worlds begin with the decision you need to rehearse.



2

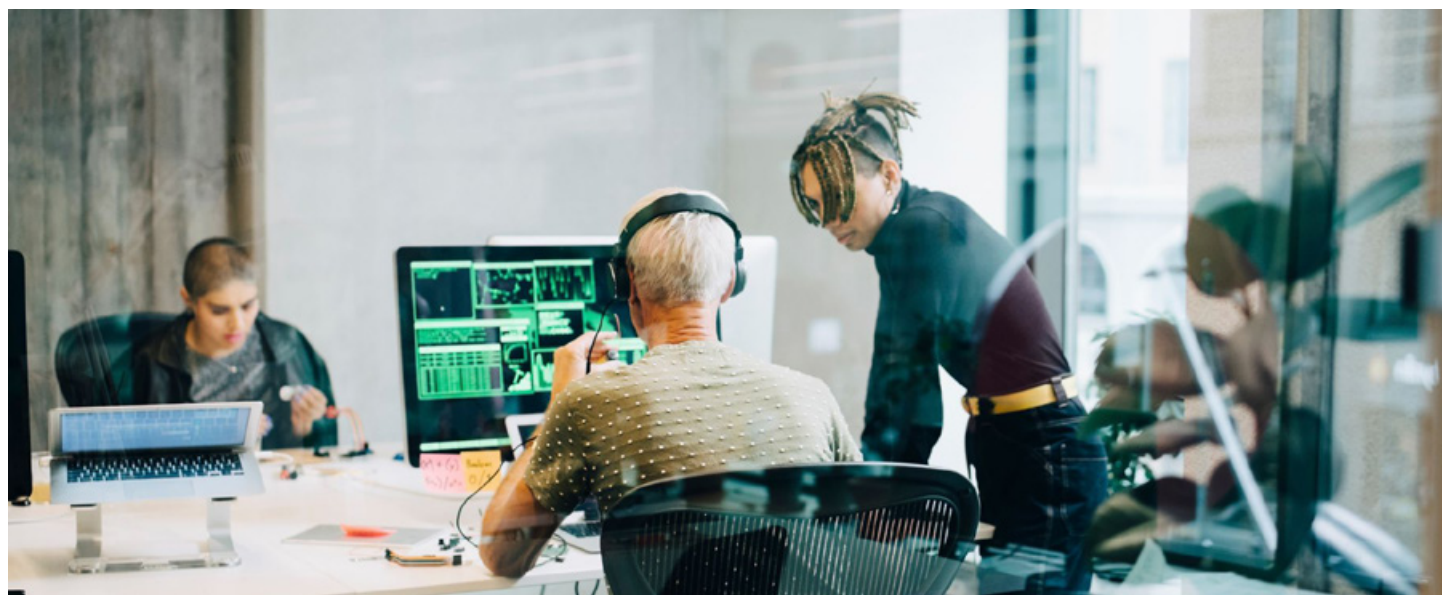
What changes when records become worlds

A synthetic world is a coherent artificial operating reality. It contains persistent entities and the relationships between them; the processes and events that move them; the business rules and constraints that shape them; and a record of what is known, assumed or still unproven.

In practice, a twin world is one way to put that idea to work: a governed environment in which a baseline reality and one or more scenario realities can be compared under explicit assumptions. It does not need to predict the future perfectly; it needs to be coherent, calibrated, reproducible and bounded.

This provides a more useful standard than simply asking whether something “looks real.” A world is realistic to the extent that it preserves the relationships, behaviours and evidence limits that matter for the decision being rehearsed. What matters is fitness for a specified use under governance rather than visual resemblance.

The idea of “known truth” also needs careful qualification. A reference world can give you known internal ground truth because the entities, events and scenario seeds are generated deliberately. This does not mean the world is true about reality; rather it means the world is inspectable, conditional on the assumptions used to build it.



3

Why AI agents need worlds to rehearse in

Agentic AI raises the stakes. A model that summarises can often be evaluated primarily through its outputs. An agent that routes cases, triggers workflows or interacts with customers, patients or suppliers must also be tested for the actions it takes, the sequence it follows, when it escalates and what happens downstream – before it goes near production.

A twin world can supply normal cases, edge cases, rare conditions, broken processes and adversarial behaviour. Because the world has known internal truth, it can score not just whether an agent reached the right answer, but how it reasoned, what it overtrusted, when it escalated and where it failed.

Credible agent sandboxes are deliberately adversarial. They include harder cases, evasive behaviour, conflicting evidence and operational stress. The point is to discover where agents, controls and policies break while the consequences are still synthetic.

Managed services is a natural near-term example. It combines repeatable operations, SLA pressure, workforce constraints and rising agent adoption. In a twin world, leaders can rehearse demand spikes, backlog growth and escalation failures before they affect live delivery.

The same architecture, different evidence boundaries

The architecture is horizontal; the evidence boundary is not. Banking, insurance, energy, healthcare, public-sector and humanitarian worlds need different real-world anchors. Same rule: build the sector kernel from the domain and label what the world can and cannot prove.

“

Realism is table stakes. Regulated organisations also need governance they can defend.

4

A decision rehearsed, not just a dataset copied

Consider a bank testing an AI-supported policy for routing customer cases. A conventional synthetic dataset can reproduce historic cases: customer attributes, transaction histories and outcomes. That is useful for privacy-safe development. For leaders, however, the broader question is what happens when the routing rule changes while capacity, customer behaviour and control obligations are also changing.

The synthetic world makes the decision under test explicit. It preserves customers, accounts, channels, case sequences, operational queues, escalation rules and control thresholds; changes one routing parameter; and compares the baseline world with the scenario world. Leaders can then assess the immediate routing

decision alongside downstream effects such as backlogs, customer-treatment exceptions, handoff failures, control breaches and situations where the AI agent begins to overtrust weak evidence.

The result is not a forecast. It is a disciplined comparison of scenarios. Each conclusion is labelled according to what the environment can support: what is synthetic, what is calibrated, what is directionally supported, what is internally validated and what still requires external evidence. That is the difference between using synthetic data to make a system plausible and using a synthetic world to make a decision reviewable.



Exhibit 1: Choose the synthetic environment the decision requires

Decision need		Better approach	Minimum evidence required	Do not use when
Test software safely	→	Seed synthetic data	Statistical similarity, privacy controls and basic utility tests.	Behavioural sequences, relationships or rare events drive the outcome.
Train or evaluate behaviour	→	Behavioural synthetic data	Temporal, relational and journey fidelity against a real or calibrated reference.	There is no credible behavioural reference or domain logic to validate against.
Explore a future state	→	Reference world	Clear assumptions, mechanism validation and sensitivity testing.	The decision requires a point forecast rather than scenario exploration.
Compare policies or operating choices	→	Twin-world rehearsal	Baseline calibration, scenario traceability and evidence labels.	The decision cannot tolerate simulation uncertainty or unvalidated assumptions.
Test autonomous agents	→	Governed agent sandbox	Known internal truth, adversarial cases, escalation tests and human override rules.	Production controls, escalation routes or accountability are not yet defined.

Exhibit 1. The point is not to climb a maturity ladder for its own sake. The right environment depends on the decision, the evidence available and the risk of being wrong.

5

The dangerous side of plausibility

“

The more realistic a synthetic asset looks, the more dangerous it becomes if no one knows where its authority ends.

Behavioural fidelity fails quietly. A learned tabular generator can match every marginal distribution you give it and still destroy the timing, sequence and entity-level dynamics that real operations actually run on. The real test is not whether the data looks right. It is whether the behaviour underneath survives.³

This is why the unit of design matters. Rows flatten the world. They may preserve attributes while losing sequence, dependency, queues, handoffs and escalation paths. Once those signals are discarded, no language model or dashboard can recover them afterward.

The risk is not merely technical. Reports published by the FCA's Synthetic Data Expert Group have highlighted questions of suitability, governance and model use in financial services.^{4,5} Technical research likewise shows that sequential synthetic data must preserve relationships across time, not merely point-level distributions.⁶ Separately, Nature research on model collapse shows how recursively training on generated data can erode distribution-tail information.⁷ The answer is not to abandon synthetic data. It is to govern it as evidence, not as decoration.

Exhibit 2: Four tests of a credible synthetic world

Test		Leadership question	Failure it prevents
Structural fidelity	⇒	Are the relevant entities, relationships, constraints and handoffs present?	A world that looks rich but omits the relationships that drive the decision.
Behavioural fidelity	⇒	Do journeys, timing, recurrence and responses behave credibly under pressure?	A world that preserves averages while destroying operational behaviour.
Decision sensitivity	⇒	Does the world respond plausibly when the decision variable changes?	A scenario engine that produces movement without meaningful cause and effect.
Evidence integrity	⇒	Can every important output be traced to a source, assumption or declared gap?	A simulation whose conclusions sound authoritative but cannot be defended.

Exhibit 2. A synthetic world is not good because it looks real. It is good when it preserves the mechanisms that matter to the decision, responds credibly when those mechanisms are disturbed, and makes unsupported conclusions visible.

6

The evidence frontier: what a synthetic world can and cannot support

The most important governance idea is the evidence frontier. Every significant output should distinguish between what is internally observable, what is weakly supported, what depends on external validation and what cannot be inferred from internal data at all.

That sounds conservative. It is actually what makes the proposition usable. Where a risk is indistinguishable from normal internal data, no amount of internal analytics will resolve it. The right response is to route the question to fieldwork, confirmation or human judgment.

In practice, this means a synthetic-world output should not simply say “low risk” or “scenario B performs better.” It should state the status of the claim: synthetic, calibrated, directionally supported, internally validated, externally confirmed or production evidence. Those labels keep simulation from masquerading as fact.

That discipline does not weaken the proposition. It is what makes the proposition credible. Analytics for the structured and observable; real-world evidence for the rest.



7

What leaders should do – and avoid

For leaders who want to move, the sequence is more important than the speed. The right starting point is not a platform choice. It is a decision that is currently unsafe, too expensive or impossible to test in the real world.

1	Do not start with a platform.	⇒ Start with a decision: the policy, workflow, operating change or agent behaviour that needs rehearsal before exposure.
2	Do not define realism as resemblance.	⇒ Define it as preservation of the mechanisms, constraints and behaviours that matter to that decision.
3	Do not build one-off datasets when a reusable world is needed.	⇒ Create sector kernels – business rules, cause-and-effect maps, controls and evidence labels – that can support multiple scenarios.
4	Do not scale a world before its evidence frontier is understood.	⇒ Every important output should carry a status label and a route to real-world validation where needed.
5	Do not deploy an agent before ambiguity has been rehearsed.	⇒ The sandbox should include failure, conflict, weak evidence and explicit escalation routes.
6	Do not confuse internal truth with external truth.	⇒ A generated world can tell you exactly what happened inside the simulation. It cannot, by itself, prove that the real world will behave the same way.

Underneath all six is one principle leaders should not delegate: treat synthetic worlds as decision infrastructure, not as a synonym for synthetic data.

8

The disciplined ambition

Synthetic data is no longer only a privacy or testing technique. It is becoming a way to rehearse decisions in settings where real-world testing is too slow, too risky, too fragmented, too sensitive or simply not yet possible. Whether a synthetic table resembles real records is no longer the decisive question. What matters now is whether a synthetic environment can help an organisation reason, experiment and govern before it acts.

Synthetic World Intelligence creates an opportunity for organisations that can combine sector knowledge, engineering discipline and assurance-grade governance. It builds

from mechanisms, preserves cross-domain coherence, supports scenario comparison, enables agent rehearsal and wraps the whole thing in evidence discipline. The strongest claim here is the disciplined one, not the inflated one. These worlds do not replace real customers, real patients, real citizens or real operational data. They make uncertainty testable, validation needs explicit and high-stakes decisions safer to rehearse before real-world exposure.

“

The ambition is simple to name and hard to earn: build worlds good enough to rehearse the decisions that matter, while staying honest about where the real evidence still has to come from.



Notes

1. World Economic Forum. "As AI blurs the lines between real and synthetic data, strong governance is essential." 15 Oct 2025. [weforum.org/stories/emerging-technologies/ai-synthetic-data-strong-governance/](https://www.weforum.org/stories/emerging-technologies/ai-synthetic-data-strong-governance/)
2. Gartner. "Gartner Announces the Top Data & Analytics Predictions." 17 Jun 2025. gartner.com/en/newsroom/press-releases/2025-06-17-gartner-announces-top-data-and-analytics-predictions
3. Yoon, J., Jarrett, D. and van der Schaar, M. "Time-series Generative Adversarial Networks." Advances in Neural Information Processing Systems 32 (2019). proceedings.neurips.cc/paper/8789-time-series-generative-adversarial-networks
4. Financial Conduct Authority / Synthetic Data Expert Group. "Report: Using Synthetic Data in Financial Services." First published 8 Mar 2024. fca.org.uk/publications/corporate-documents/report-using-synthetic-data-financial-services
5. Financial Conduct Authority / Synthetic Data Expert Group. "Generating and using synthetic data for models in financial services: governance considerations." 19 Aug 2025. fca.org.uk/publications/corporate-documents/synthetic-data-models-financial-services-governance-considerations
6. Yoon, J., Jarrett, D. and van der Schaar, M. "Time-series Generative Adversarial Networks." Advances in Neural Information Processing Systems 32 (2019). proceedings.neurips.cc/paper/8789-time-series-generative-adversarial-networks
7. Shumailov, I., et al. "AI models collapse when trained on recursively generated data." Nature 631, 755-759 (2024). doi.org/10.1038/s41586-024-07566-y

Further reading

Jordon, J., et al. "Synthetic Data – what, why and how?" Royal Society / Alan Turing Institute explainer, 2022. doi.org/10.48550/arXiv.2205.03257

European Data Protection Supervisor. "Synthetic Data – TechSonar." edps.europa.eu/press-publications/publications/techsonar/synthetic-data_en

MIT Data to AI Lab. "The Synthetic Data Vault." 2016. hdi-dai.lids.mit.edu/projects/synthetic-data-vault/

The views reflected in this article are the views of the author and do not necessarily reflect the views of the global EY organization or its member firms.



Authors



Dr. Kostis Chlouverakis

Partner, Consulting Services –
Europe Central Financial Services
AI Leader, EY Greece



Richard Clough

Partner, EY Global Chief Data Officer,
Growth & Innovation, EY LLP

EY | Building a better working world

EY is building a better working world by creating new value for clients, people, society and the planet, while building trust in capital markets.

Enabled by data, AI and advanced technology, EY teams help clients shape the future with confidence and develop answers for the most pressing issues of today and tomorrow.

EY teams work across a full spectrum of services in assurance, consulting, tax, strategy and transactions. Fueled by sector insights, a globally connected, multidisciplinary network and diverse ecosystem partners, EY teams can provide services in more than 150 countries and territories.

All in to shape the future with confidence.

EY refers to the global organization, and may refer to one or more, of the member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients. Information about how EY collects and uses personal data and a description of the rights individuals have under data protection legislation are available via ey.com/privacy. EY member firms do not practice law where prohibited by local laws. For more information about our organization, please visit ey.com.

© 2026 EYGM Limited.
All Rights Reserved.

BMC Agency
GA 114944545

SCORE 115549-26-GBL
ED None

This material has been prepared for general informational purposes only and is not intended to be relied upon as accounting, tax, legal or other professional advice. Please refer to your advisors for specific advice.

ey.com