





## Document overview

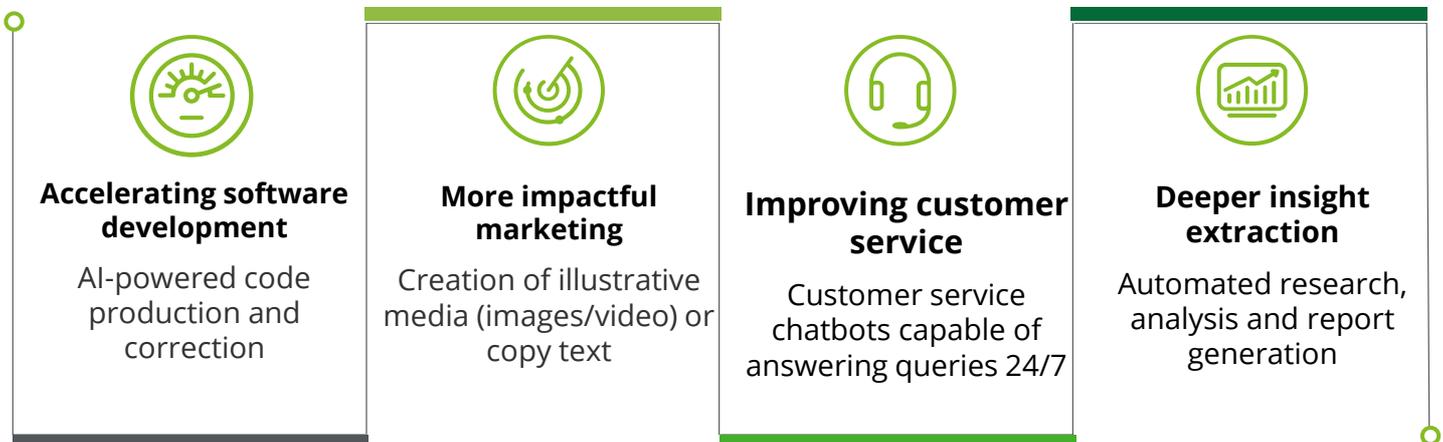
This Point of View (PoV) showcases Deloitte’s perspective on the role of data governance in generative artificial Intelligence (AI). We demonstrate how organisations can only unlock the full potential value from their Generative AI use cases through robust data governance capabilities. Data governance is also an essential component of safeguarding against the risks presented by this new, fast evolving technology.

“ In the [Summer 2023 Fortune Deloitte CEO Survey](#), **55% of CEOs indicated that they are evaluating & experimenting with Generative AI**, and **37% of them are currently implementing Generative AI** to some degree. ”

Organisations are moving fast to employ Generative AI models in order to transform their business, however they are often failing to consider the **importance of established data governance capabilities** which are required to fully **support and extract the maximum value** from Generative AI. **Without robust data governance capabilities**, the **potential impact and value added by Generative AI will be severely limited** and may even expose organisations to **data and cybersecurity risks**.

## What is Generative AI?

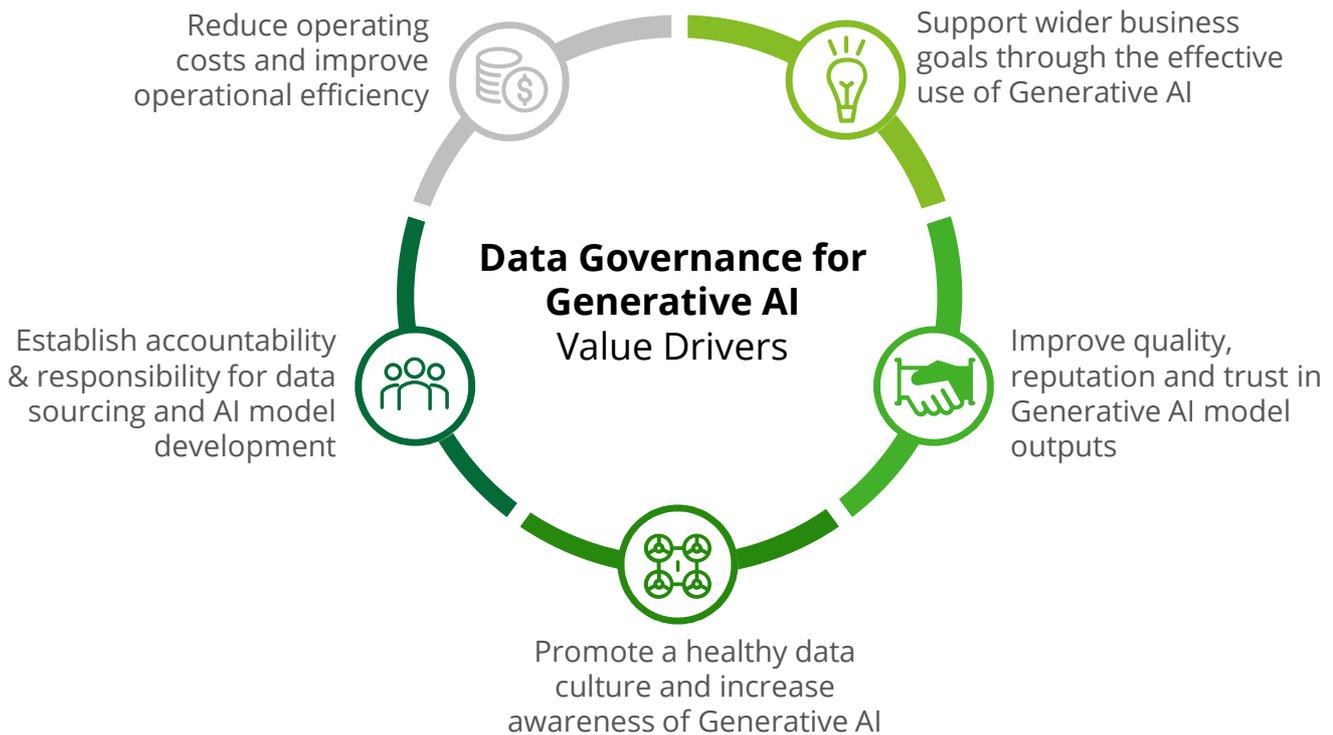
Generative AI is a subset of AI that can generate text, images, audio, video. Generative AI uses machine learning models to identify patterns within underlying existing data (known as ‘training data’), which is then used to produce new content in response to external prompts inputted by users. Generative AI has the potential to reduce the marginal cost of generating knowledge-intensive content significantly. It has the potential to facilitate the creation of new ideas and information at a faster rate, and potentially of higher quality, than what could be achieved through human effort alone.



Generative AI is proving to be a disruptive and fast-changing technology. It is increasingly clear that many organisations must move to leverage this technology if they wish to stay competitive in the market.

# Why is data governance important in Generative AI?

Data governance plays a pivotal role in fostering innovation in the evolving AI landscape by ensuring responsible data practices, mitigating biases, and safeguarding privacy. A robust data governance strategy is the key to unlocking the full potential of your Generative AI use cases.



## Our Perspective

1

### **Generative AI is only as good as your data quality**

Data governance frameworks ensure the highest standards of data quality and integrity, so that models are properly governed and outputs are well-understood

2

### **Empowering Generative AI through data literacy**

Organisations who maintain a culture of data literacy are more successful in consistently unlocking value from their Generative AI platforms

3

### **Using Generative AI responsibly**

Data governance plays a critical role in promoting secure, lawful, ethical and transparent AI driven decision-making

# Fuelling the future: Generative AI is only as good as the data

Generative AI models must be trained on data that is fit for purpose to achieve accurate and quality output. Explainable outcomes are essential to instil trust in Generative AI systems, enabling users to comprehend the reasoning behind the outcomes they produce. Effective data governance practices are necessary to ensure quality, integrity, and representativeness of the data.

## Data governance considerations



Implement data governance practices that ensure **training data** used in foundation models is **accurate, reliable, and high quality**



Ensure **clear visibility of data lineage** to **track the origin and transformation** of the data used to train the AI model



Ensure sufficient **data access controls** are in place for input/training data depending on the level of sensitivity, and in line with GDPR compliance for personal data relating to data subjects



Well **defined processes and documentation** promotes transparency and helps stakeholders understand the data required for their Generative AI models



Adapt the current **operating model** to promote **AI governance, decision making, and cross-domain collaboration**

## Implications for Generative AI



Ensures input training **data is fit-for-purpose** leading to **higher quality outputs and insights**



Allows users to **identify potential sources of bias, errors, and unethical content** that may be present in generated outputs



**Mitigates security risks** and ensures that only **approved users have access** to Generative AI capabilities and data



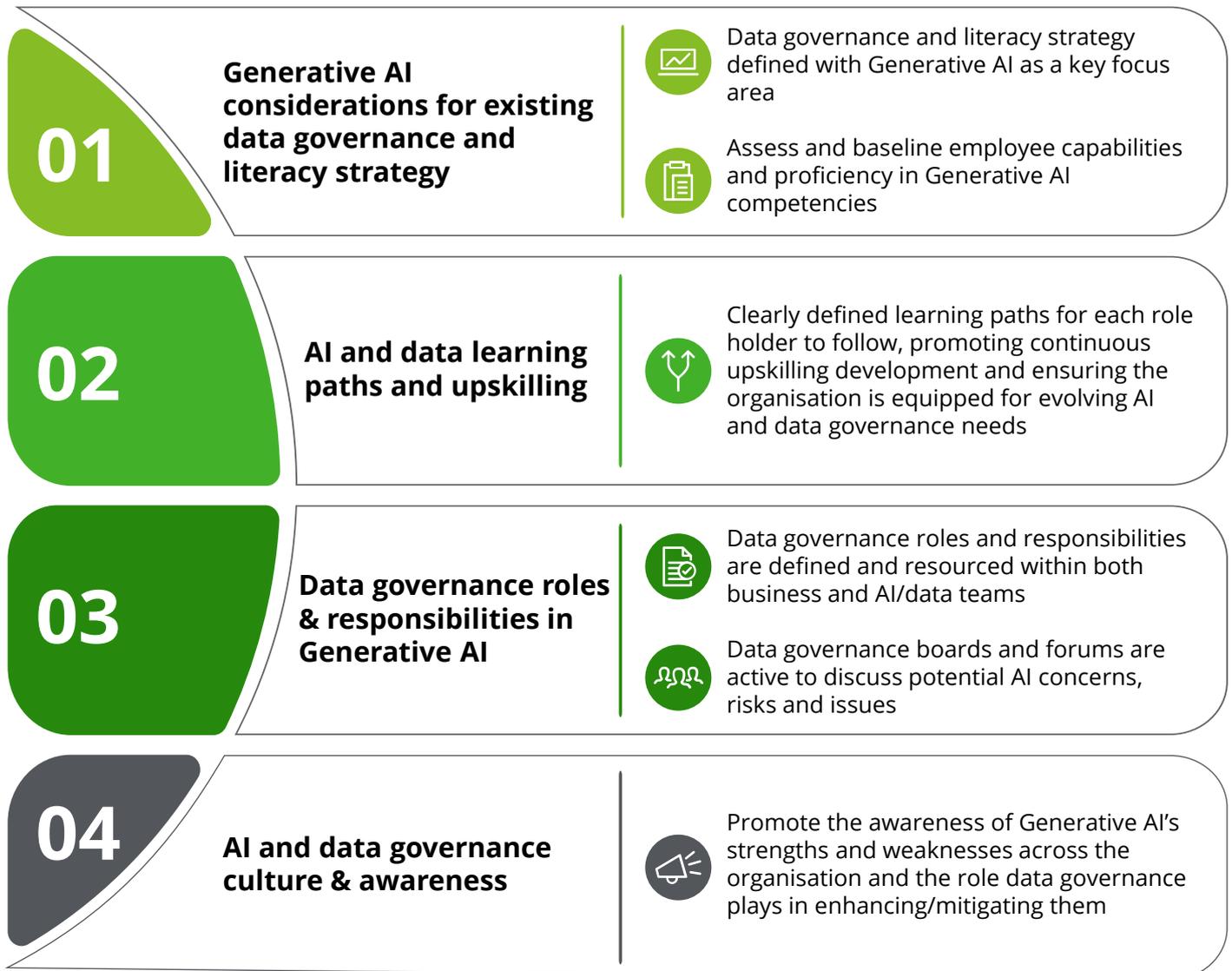
Promotes **transparency** and helps stakeholders **understand the data required** for their Generative AI models



Ensures roles and responsibilities are defined, with **clear accountability for the proper use** of Generative AI across the organisation

# Empowering Generative AI through data literacy

Organisations that invest in upskilling and developing the data literacy capabilities of their workforce are more successful in consistently unlocking value from their Generative AI platforms.



# Using Generative AI responsibly

The fast-evolving landscape of Generative AI presents a massive opportunity for organisations to revolutionise their business and build on new opportunities, but this comes with the obligation to ensure that Generative AI is used in a responsible and ethical manner that minimises risk to organisations and individuals.

	Why it matters	What you need to do
 <p><b>The dangers of bias</b></p>	<p>Generative AI can perpetuate <b>harmful and untrue stereotypes</b> if these are reflected in the underlying training data, known as <b>'biased data'</b></p>	<p>Consider <b>where you are sourcing your data from</b>, what biases may impact your Generative AI models and how to detect and mitigate existing bias in foundation models</p>
 <p><b>'Adversarial AI'</b></p>	<p>Generative AI models can be <b>abused by malicious actors to do harm</b> to organisations or individuals</p>	<p><b>IT security protocols and risk management frameworks</b> with mitigation plans can help protect against the malicious use of Generative AI</p>
 <p><b>Hallucination</b></p>	<p>Generative AI models that <b>assert erroneous or wrong conclusions</b> suffer from a concept known as 'hallucination'</p>	<p>Data quality <b>monitoring and data quality issue remediation</b> processes ensure only <b>data that is fit for purpose</b> is used to train Generative AI, providing <b>better assurance for model outputs.</b></p>
 <p><b>Emergent properties</b></p>	<p><b>Predicting and understanding why certain outputs are produced can be difficult</b> with open-source Generative AI models</p>	<p><b>Establishing data standards</b> for the data used for training and testing Generative AI models <b>reduces the likelihood of emergent properties</b> caused by incorrect or incomplete data</p>
 <p><b>Fair use and intellectual property</b></p>	<p>External training data procured via third parties or mass web-scraping <b>raises intellectual property/copyright risks</b></p>	<p><b>Use of data catalogues</b> to understand where you are sourcing your data from and <b>outlining ownership of data</b> ensures that data is sourced from approved sources and used in line with data owner expectations</p>
 <p><b>Data privacy and the use of personal data</b></p>	<p>The <b>use of personal data</b> such as names, images of people/faces <b>poses significant security, regulatory and legal considerations</b></p>	<p>Use of information classification alongside retention and deletion schedules can be used to <b>effectively manage personal data in a lawful and compliant manner</b></p>

# Key takeaways

Implementing effective data governance that is capable of supporting Generative AI can be difficult. Whilst many organisations lack effective data governance capabilities, there is still opportunity for your organisation to consider how data governance should support your Generative AI strategy.

## Organisations that want to make best use of Generative AI will benefit from having mature data governance capabilities



- Training data that is **well governed** will produce higher quality outputs for your Generative AI models
- Higher quality outputs enable **extraction of greater value** from your use cases, through gaining **deeper insights** or enabling more potential functionality of your use case
- Data governance ensures that your organisation's **operating model is adapted to accommodate Generative AI** including defined roles, responsibilities, and learning paths

## Without data governance, Generative AI capabilities are limited and will expose organisations to a number of risks



- Data governance ensures that organisations **comply with relevant data protection regulations** when using Generative AI
- Data governance **promotes ethical and responsible AI practices**, including addressing biases and privacy concerns with training data
- Data governance instils confidence in stakeholders and users by ensuring that the **data used is trustworthy, secure, reliable, and aligned to the organisation's values and goals**

## Data governance should be continuously implemented throughout Generative AI as part of a wider data culture of data



- Generative AI is a disruptive technology which can be difficult to embed. Data governance will **enable smoother adoption of your Generative AI use cases**
- Data governance fosters a **collaborative and data-centric culture**, reinforcing the idea that data and information are valuable strategic assets
- Improving individual awareness of data governance and Generative AI will **improve employee competency and efficiency** at scale

## Start a conversation with us

---



### Ray Pulapaka

Head of Strategy, Governance & Architecture  
[rapulapaka@deloitte.co.uk](mailto:rapulapaka@deloitte.co.uk)



### Elliot Stamp

Head of Data Governance & Operations  
[estamp@deloitte.co.uk](mailto:estamp@deloitte.co.uk)



### Mo El-Jarad

Manager - Data Governance & Operations  
[meljarad@deloitte.co.uk](mailto:meljarad@deloitte.co.uk)

This publication has been written in general terms and we recommend that you obtain professional advice before acting or refraining from action on any of the contents of this publication. Deloitte LLP accepts no liability for any loss occasioned to any person acting or refraining from action as a result of any material in this publication.

Deloitte LLP is a limited liability partnership registered in England and Wales with registered number OC303675 and its registered office at 1 New Street Square, London EC4A 3HQ, United Kingdom.

Deloitte LLP is the United Kingdom affiliate of Deloitte NSE LLP, a member firm of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"). DTTL and each of its member firms are legally separate and independent entities. DTTL and Deloitte NSE LLP do not provide services to clients. [Please click here to learn more about our global network of member firms.](#)