

POLITECNICO MILANO 1863 | DIPARTIMENTO DI INGEGNERIA GESTIONALE

Emerging Issues: Artificial Intelligence and Public Sector Financial Reporting

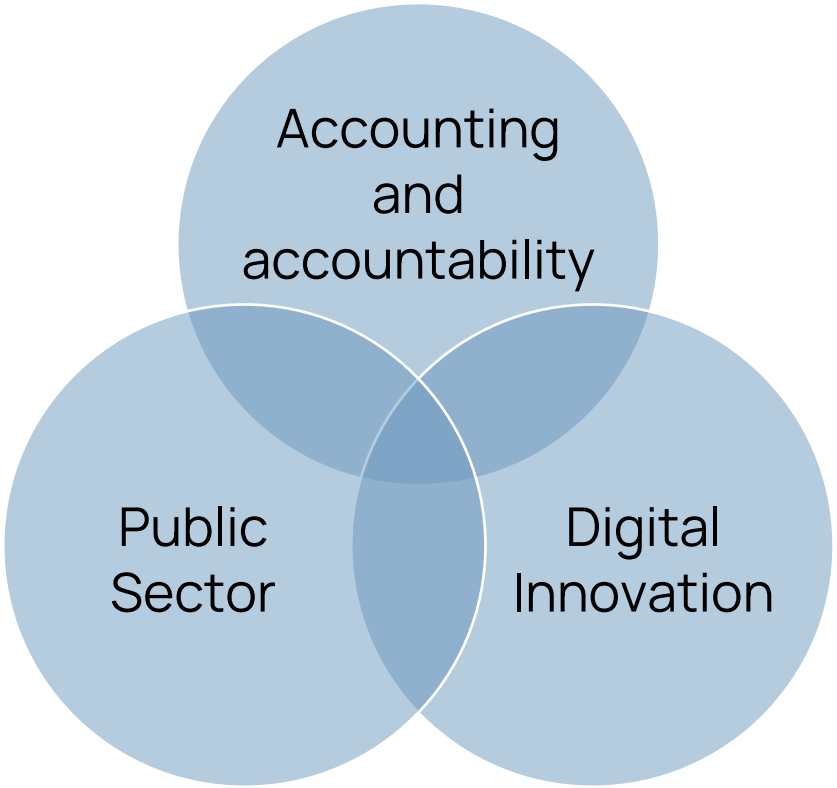
Governance, accountability, and practical implications

12 March 2026

Deborah Agostino

Deborah Agostino

Associate Professor in Accounting, Finance and Control @Politecnico di Milano
Adjunct Professor @ City University of Hong Kong



Editorial Activities

- Associate Editor of Financial Accountability and Management
- Associate Editor of European Management Journal (Public Management Field)
- Editorial board member of Public Money and Management
- Editorial advisory board member of Accounting, Auditing and Accountability Journal

Engagement with the research community

- *Board Member, International Research Society of Public Management (IRSPM)*
- *Board Member, Public Service Accounting and Accountability Group (PSAAG)*
- *Scientific Committee Member, EIASM Public Sector Accounting and Accountability Conference*

Agenda

01

Introduction to the AI

02

Deeper dive into AI models

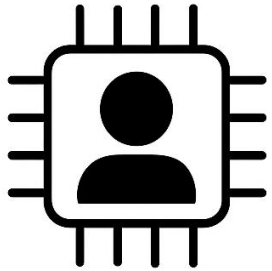
03

AI implications for accountability and accounting

04

Key Questions for IPSASB

From AI to machine learning, generative AI, and agentic systems



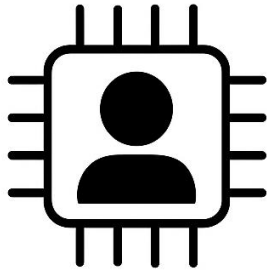
ARTIFICIAL INTELLIGENCE

Since the 1950s

Umbrella term for systems that perform tasks associated with human intelligence

AUTOMATION

AI evolution: from AI to Agentic AI



ARTIFICIAL INTELLIGENCE

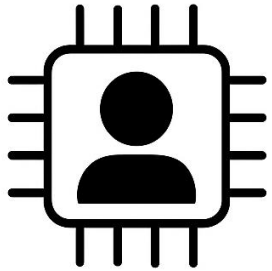
Since the 1950s

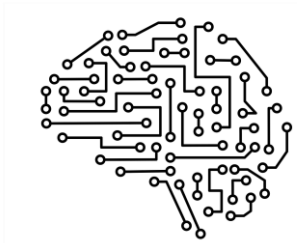
Umbrella term for systems that perform tasks associated with human intelligence

AUTOMATION

- **What it mainly does**
  - Executes repetitive, rule-based, structured tasks
- **Potential use in accounting and reporting**
  - transaction processing
  - reconciliations
  - invoice matching
  - closing routines
  - control execution

AI evolution: from AI to Agentic AI





ARTIFICIAL INTELLIGENCE

Since the 1950s

Umbrella term for systems that perform tasks associated with human intelligence

AUTOMATION

MACHINE LEARNING

Mainstream adoption accelerated in the 2010s

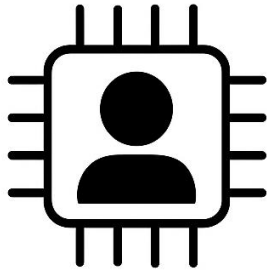
Focus on pattern recognition and prediction from data

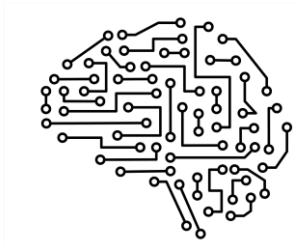
PREDICTION

- **What it mainly does**
  - Identify patterns
  - classify cases
  - forecast outcomes
  - detect anomalies

- **Potential usage in accounting and reporting**
  - fraud and anomaly detection
  - risk scoring
  - revenue and expenditure forecasting
  - exception monitoring

AI evolution: from AI to Agentic AI







ARTIFICIAL INTELLIGENCE

Since the 1950s

Umbrella term for systems that perform tasks associated with human intelligence

AUTOMATION

MACHINE LEARNING

Mainstream adoption accelerated in the 2010s

Focus on pattern recognition and prediction from data

PREDICTION

GENERATIVE AI

Public breakout in 2022-23

Focus on content generation with foundation models and LLMs

CREATION

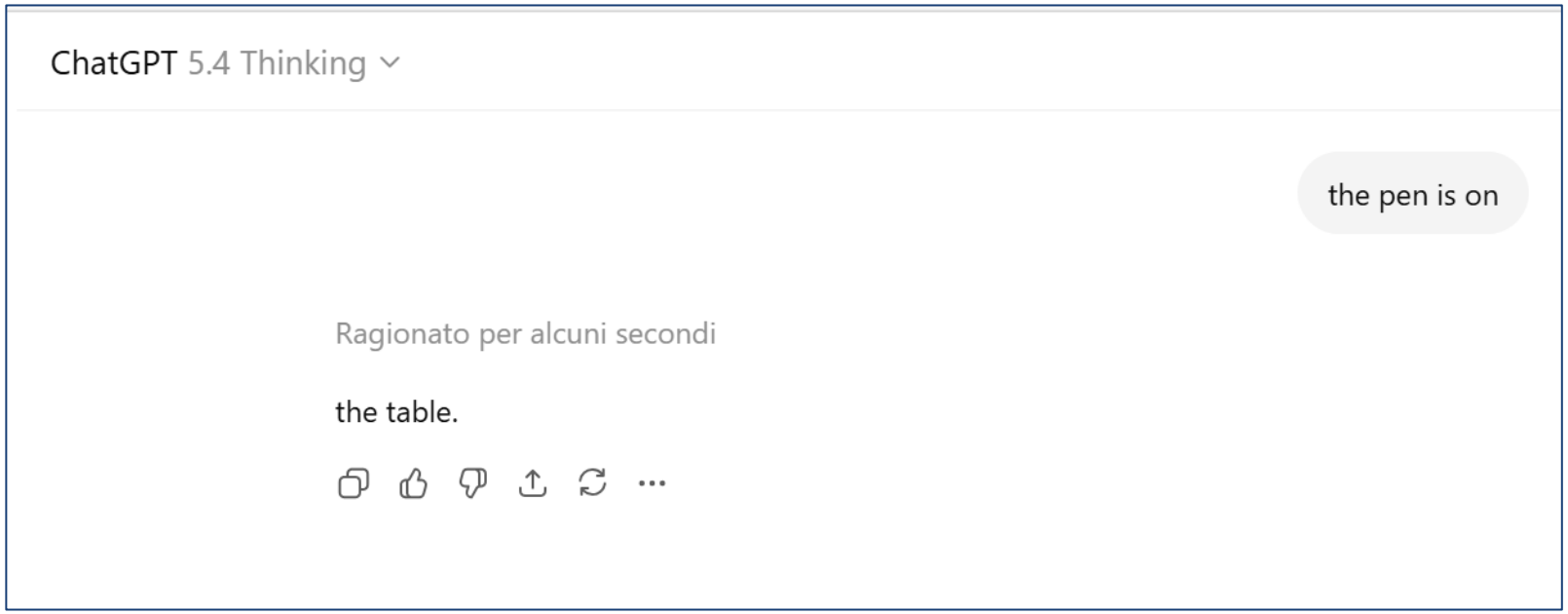
- **What it mainly does**
  - Generate text, summaries, code, synthetic content

- **Potential usage in accounting and reporting**
  - memo drafting
  - disclosure drafting
  - summarising standards
  - management commentary support
  - note preparation

LLM (Large Language Model) in simple terms

- An LLM does not “understand” in a human sense; it predicts the next token given a context.
- A **token** is the basic unit of text that an LLM processes: a piece of a word, a whole word, or a symbol, produced by a tokenizer so that text can be converted into numbers the model can work with.
- “The teacher praised the student” → ["The", " teacher", " praised", " the", " student"]
- Common words tend to be single tokens.
- Position is fundamental: The teacher praised the student =/= The student praised the teacher

LLM in simple terms: next token prediction



- An LLM **assigns probabilities** to possible next token
- Depending on the settings, it selects one of the most likely tokens.
- By repeating this process **one word at a time**, it generates sentences, paragraphs, and longer outputs..

LLM in simple terms: AI training

- During training, the model processes very large volumes of text and other data, depending on the model.
- It learns statistical regularities in language, style, and associations between concepts.
- It stores patterns in parameters rather than as a structured encyclopedia.
- Without grounding or verification, it can produce plausible but incorrect outputs.

It reproduces learned patterns

Remember that:

- Fluency is not evidence of accuracy: it cannot distinguish true from false. Answers based on “statistical intuition”.
- Factual validation must come from authoritative sources.

LLM in simple terms: AI parameters

Large LLMs (e.g. GPT-5) have billions or trillions of parameters

A parameter is an internal numerical weight of the model, learned during training.
These weights are stored in matrices and help the model recognize patterns in language and predict likely outputs.

Parameters represent the size and capacity of the model.
More parameters allow the model to capture more complex patterns, but they also require more data, more computing power, and more energy to train and run.

Example:
If the model reads:
"The doctor examined the ..."
its parameters help it assign higher probability to words like:
patient, child, or woman
than to words like:
mountain or airplane.

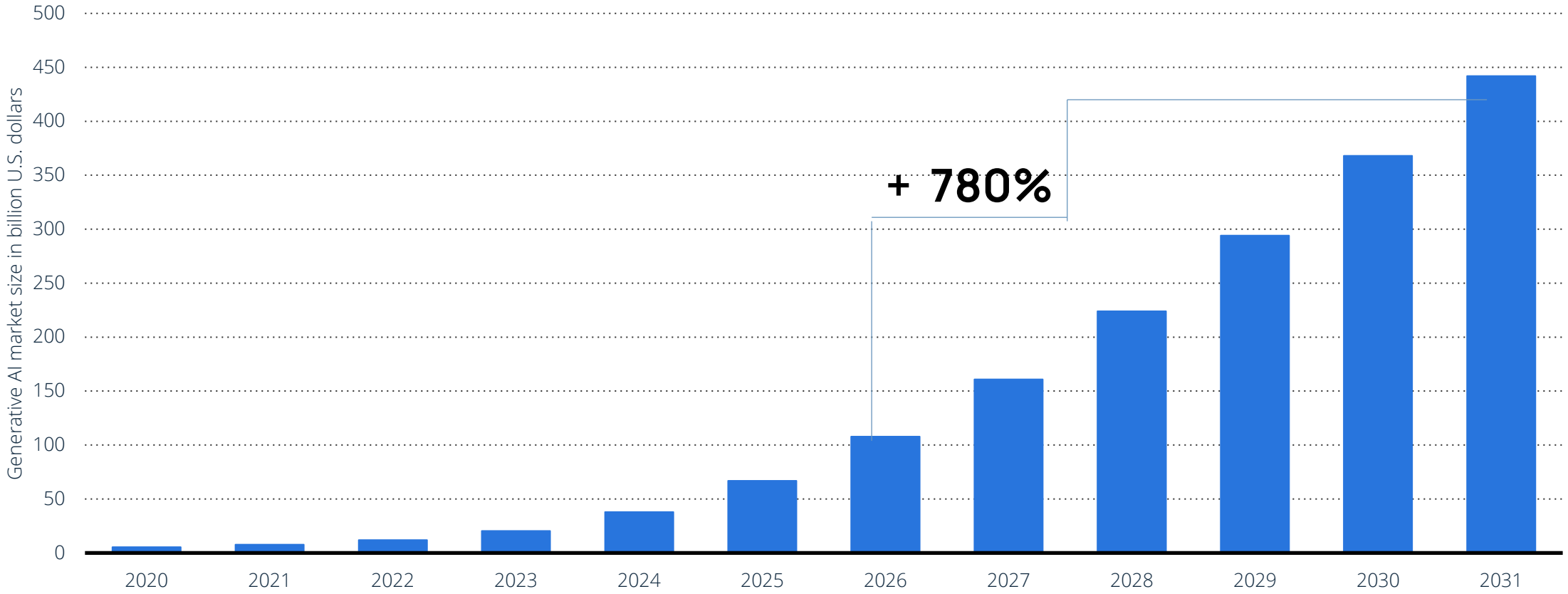
LLM: strenghts and limitations

Strenghts	Limitations
Process and summarise large volumes of text quickly	May produce plausible but inaccurate answers (“hallucinations”)
Draft memos, reports, and explanations	May reflect bias present in data or design
Make complex information more accessible	Often lack transparency in how outputs are generated
Support productivity in language-based tasks	Require human review in high-accountability contexts

LLMs can significantly support human work, but they do not replace professional judgement.

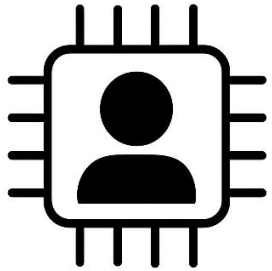
The expected growth of Generative AI

Generative AI market size worldwide 2020-2031 in billion US dollars



Note(s): Worldwide; 2020 to 2031
Further information regarding this statistic can be found on page 8.
Source(s): Statista; ID 1449838

AI evolution: from AI to Agentic AI

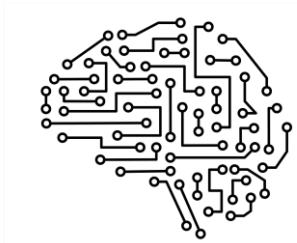


ARTIFICIAL INTELLIGENCEz

Since the 1950s

Umbrella term for systems that perform tasks associated with human intelligence

AUTOMATION



MACHINE LEARNING

Mainstream adoption accelerated in the 2010s

Focus on pattern recognition and prediction from data

PREDICTION



GENERATIVE AI

Public breakout in 2022-23

Focus on content generation with foundation models and LLMs

CREATION



AGENTIC AI

Emerging pattern

Focus on multi-step systems that can plan, use tools, access memory, and trigger actions.

ACTION

Agentic AI system: example

- A user asks: “Prepare a short briefing on the main changes in this reporting standard and compare them with last year’s guidance.”

- A stand-alone LLM may generate an answer from the prompt.

- An agentic AI system could instead:
  - find the relevant standard,
  - retrieve last year’s guidance,
  - compare the two documents,
  - summarise the main differences,
  - prepare a draft briefing,
  - route the draft for human review

Traditional LLM vs Agentic system

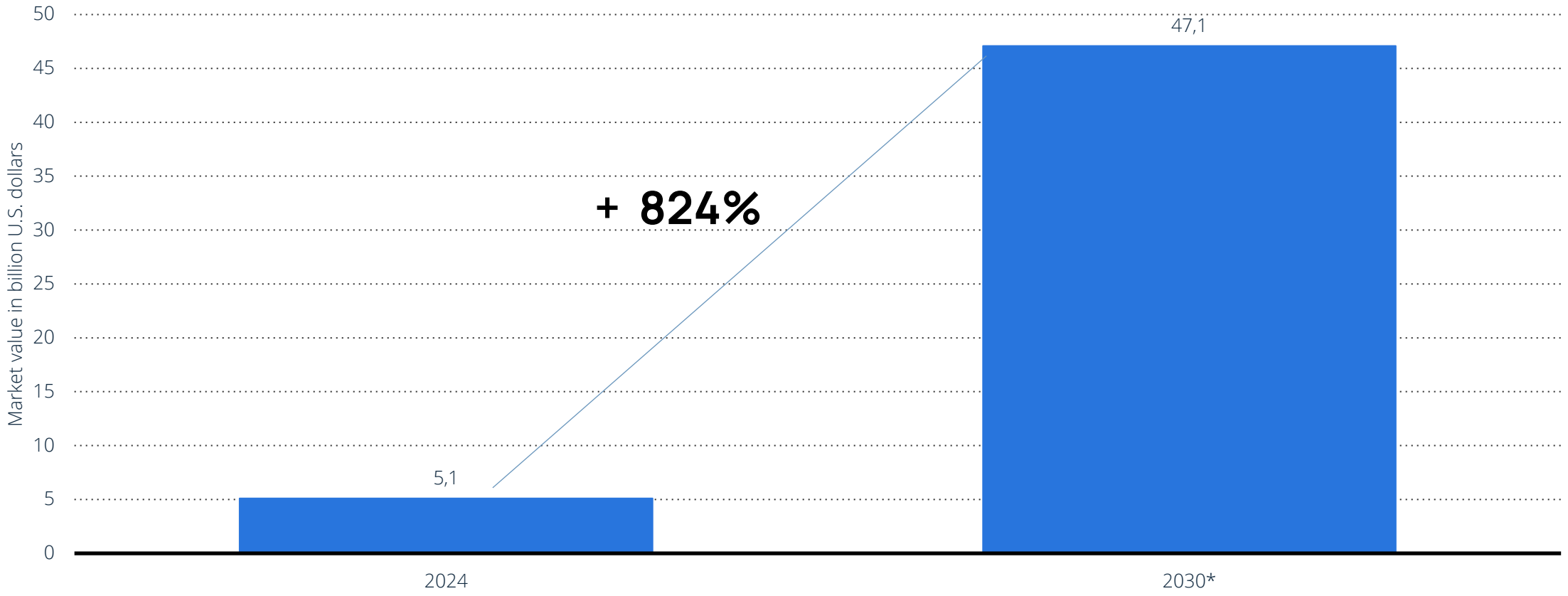
Traditional LLM	Agentic AI
Answers prompts	Pursues goals
Generates text	Generates text and coordinates actions
Usually single-step	Usually multi-step
Works mainly inside the conversation	Can interact with external tools and systems
Supports human work	Can partially execute workflows
Main risk: inaccurate or invented content	Main risk: incorrect action, weak oversight, unclear responsibility

Why agentic systems raise new governance issues

- They can retrieve data, call tools, and trigger actions.
- Errors can propagate beyond text into workflows.
- Risk rises when the system affects records, approvals, or external communication.
- Human approval, logging, and override mechanisms become critical.

The expected growth of Agentic AI

Global market value of agentic AI in 2024 & 2030 (in billion US dollars)



Source(s): Capgemini; ID 1552183

Global market value of agentic AI 2030| Statista

Agenda

01

Introduction to the AI

02

Deeper dive into AI models

03

AI implications for accountability and accounting

04

Key Questions for IPSASB

The type of AI tool also makes the difference

	Public SaaS / consumer AI
Example	ChatGPT, Gemini, Claude
How it functions	You write a prompt. The data leaves your computer, travels over the internet, reaches the provider's servers (e.g., OpenAI in the US), and goes into their "black box." Data handling: By default, the provider can see the data and often use it to train future models.
Advantages	Maximum capability (you use the smartest models), fast access, low or zero cost, no setup, strong general-purpose capability
Main Challenges	**PRIVACY** – Sensitive data may leave your controlled environment. **DATA CONTROL** – Limited control over retention, logging, and where data is processed. **COMPLIANCE** – Extra care is needed for confidential or regulated information.

The type of AI tool also makes the difference

	Public SaaS / consumer AI	Self-hosted / on-premises open-weight models
Example	ChatGPT, Gemini, Claude	Llama, Mistral, Qwen e DeepSeek
How it functions	You write a prompt. The data leaves your computer, travels over the internet, reaches the provider's servers (e.g., OpenAI in the US), and goes into their "black box." Data handling: By default, the provider can see the data and often use it to train future models.	The model runs on organisation-controlled devices or infrastructure. Data can remain inside the organisation's controlled environment.
Advantages	Maximum capability (you use the smartest models), fast access, low or zero cost, no setup, strong general-purpose capability	Greater control over data, security, and customisation Supports sovereignty requirements Can be integrated with internal systems
Risks	**PRIVACY** – Sensitive data may leave your controlled environment. **DATA CONTROL** – Limited control over retention, logging, and where data is processed. **COMPLIANCE** – Extra care is needed for confidential or regulated information.	**COST** – You need hardware, setup, and maintenance. **ICT LOAD / RESPONSIBILITY** – Security and model management stay with you. **MODEL LIMITS** – Local models may be weaker on complex tasks.

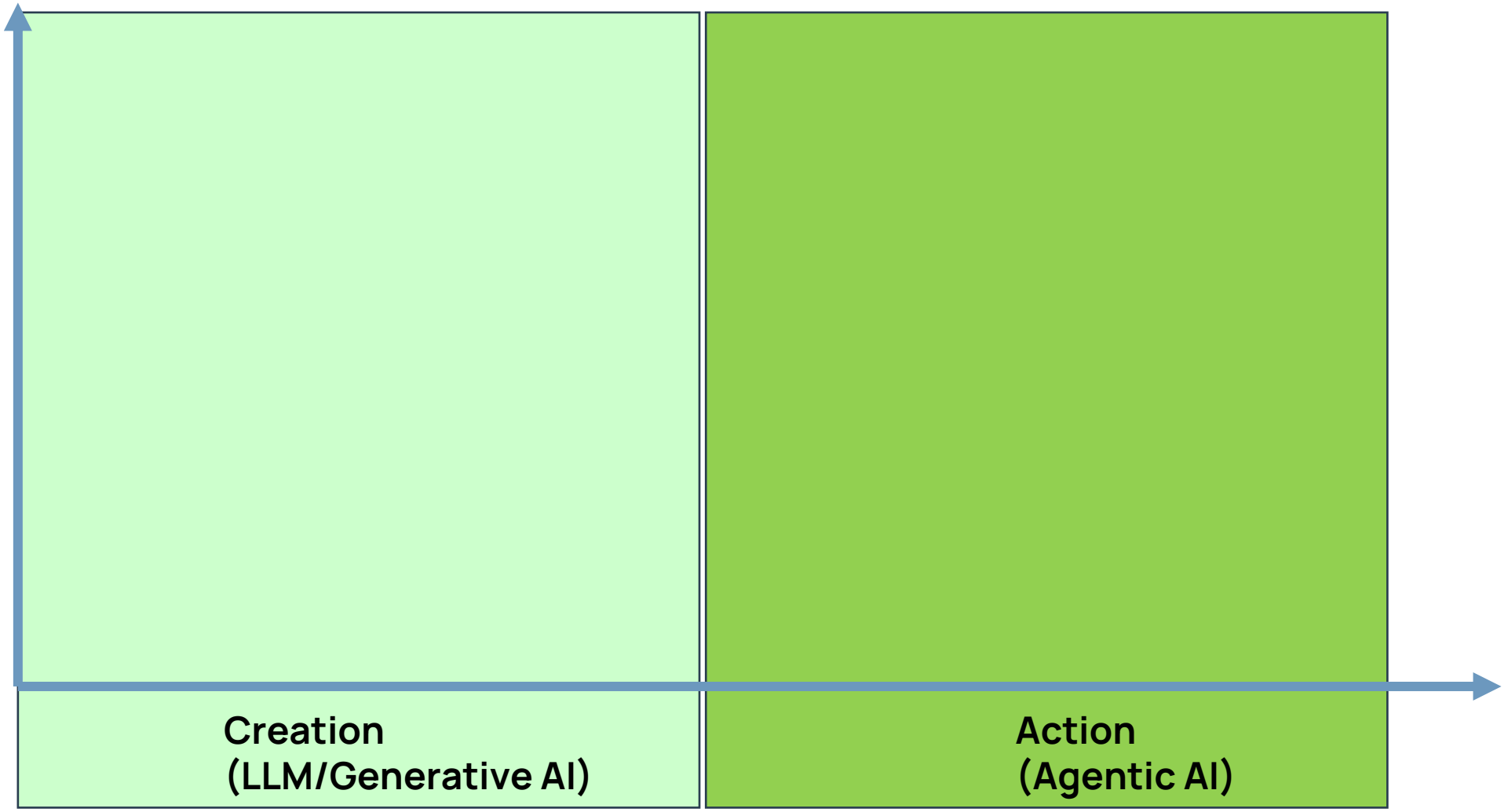
The type of AI tool also makes the difference

	Public SaaS / consumer AI	Self-hosted / on-premises open-weight models	Managed cloud / enterprise API
Example	ChatGPT, Gemini, Claude.ai	Llama, Mistral, Qwen e DeepSeek	Azure OpenAI, Amazon Bedrock, or Google Vertex AI.
How it functions	You write a prompt. The data leaves your computer, travels over the internet, reaches the provider's servers (e.g., OpenAI in the US), and goes into their "black box." Data handling: By default, the provider can see the data and often use it to train future models.	The model runs on organisation-controlled devices or infrastructure. Data can remain inside the organisation's controlled environment.	The organisation accesses advanced models through enterprise cloud services under commercial terms and technical controls. Governance depends on contract terms, region, configuration, and monitoring. (Based on contratcs your data could be hosted and managed in your "secure" space (your tenant))
Advantages	Maximum capability (you use the smartest models), fast access, low or zero cost, no setup, strong general-purpose capability	Greater control over data, security, and customisation Supports sovereignty requirements Can be integrated with internal systems	Strong capability with better enterprise governance Easier access control, logging, and integration More suitable for controlled organisational use
Risks	**PRIVACY** – Sensitive data may leave your controlled environment. **DATA CONTROL** – Limited control over retention, logging, and where data is processed.	**COST** – You need hardware, setup, and maintenance. **ICT LOAD / RESPONSIBILITY** – Security and model management stay with you. **MODEL LIMITS** – Local models may be	**LOCK-IN** – You may become tied to one provider. **COST** – Enterprise use can become expensive at scale. **CONTROL** – Governance still depends on

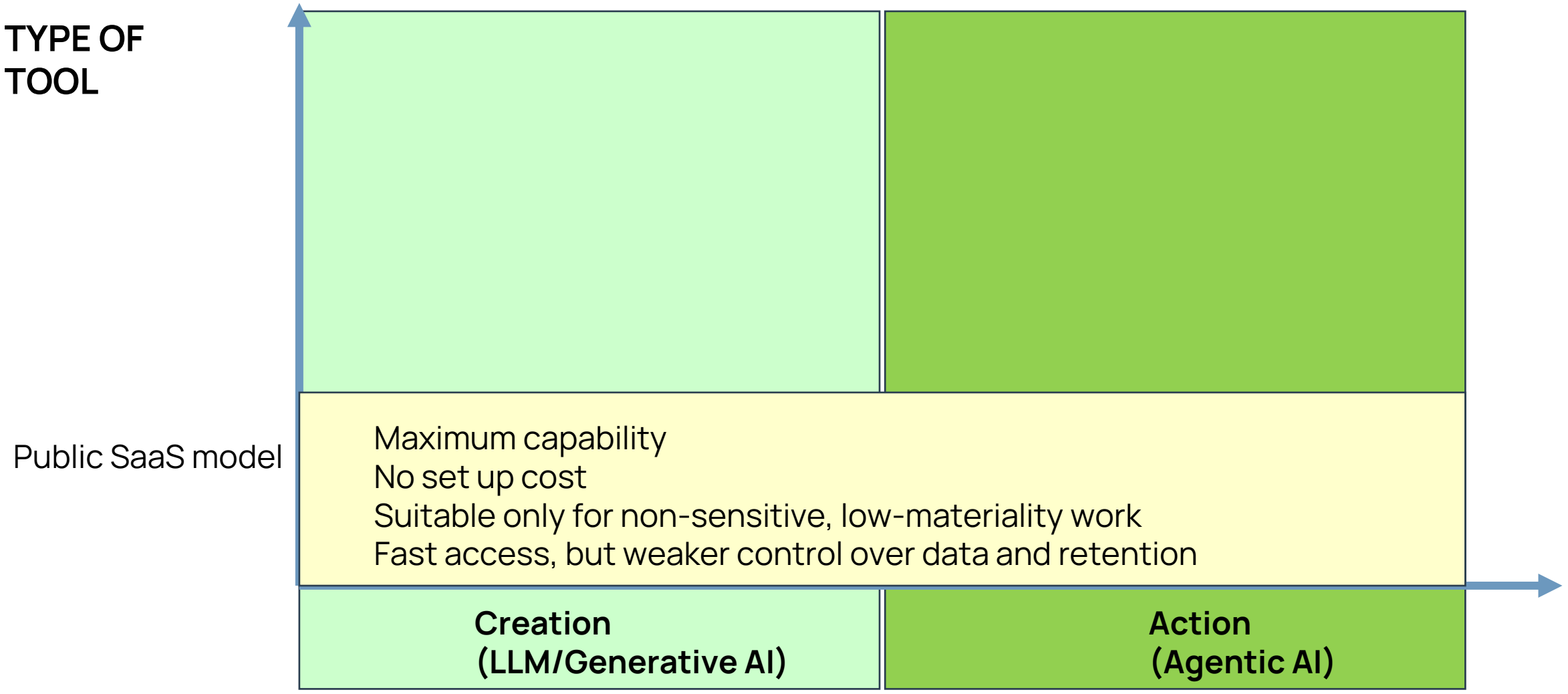
How to choose between the different models?

Selection criterion	Guiding question
Data sensitivity	*Can this data safely leave the organization's protected environment?*
Required capability	*Do we need the highest possible performance, or is a smaller/local model sufficient for the task?*
Privacy and compliance	*Is this option compatible with our legal, regulatory, and data protection obligations?*
Control and auditability	*Can we document, monitor, and review how the tool is used and how outputs are produced?*
Cost and infrastructure	*Is this option sustainable in terms of setup, maintenance, and operating costs?*
Internal capabilities	*Do we have the technical skills and organizational maturity to manage this solution?*
Vendor dependence / lock-in	*How much dependence on an external provider are we willing to accept?*

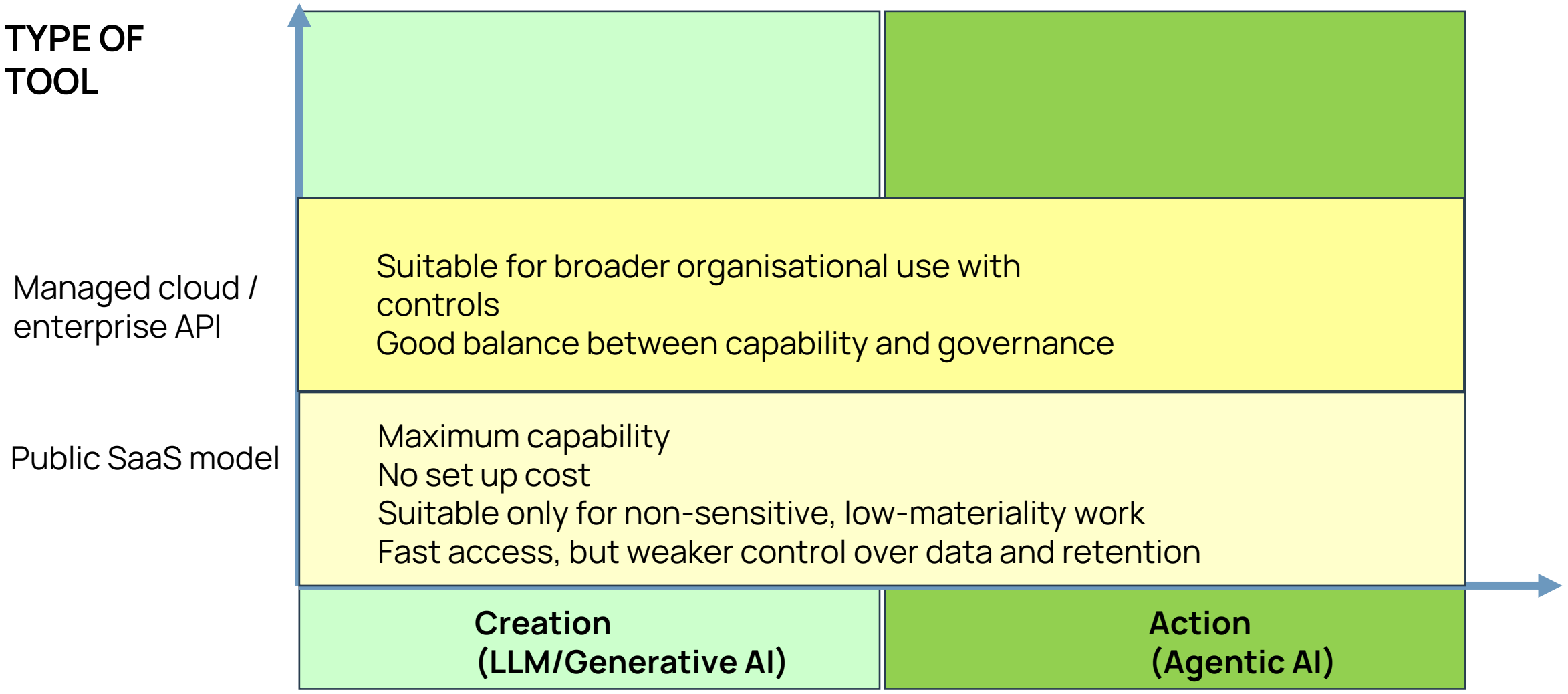
Two different choices



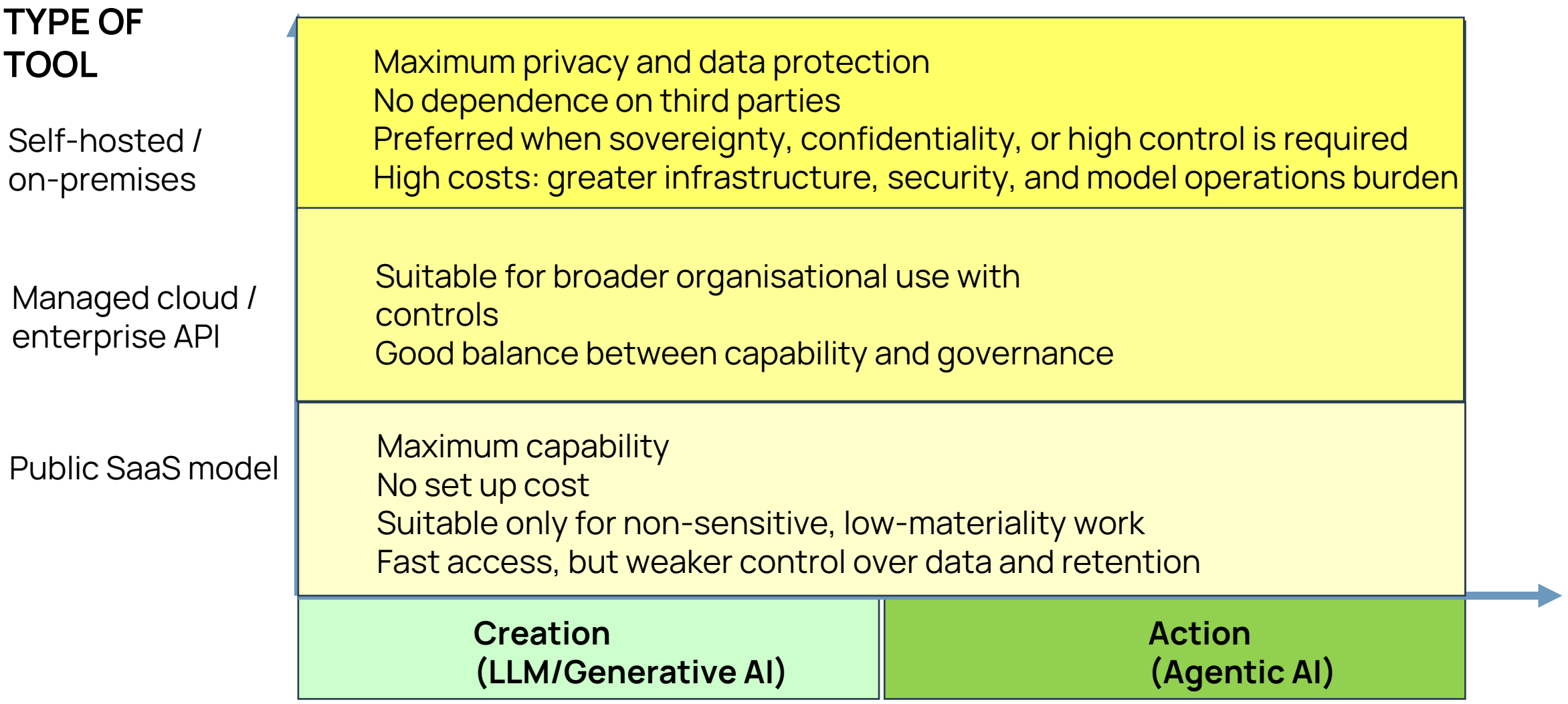
Two different choices



Two different choices



Two different choices



Agenda

01

Introduction to AI

02

How AI models work

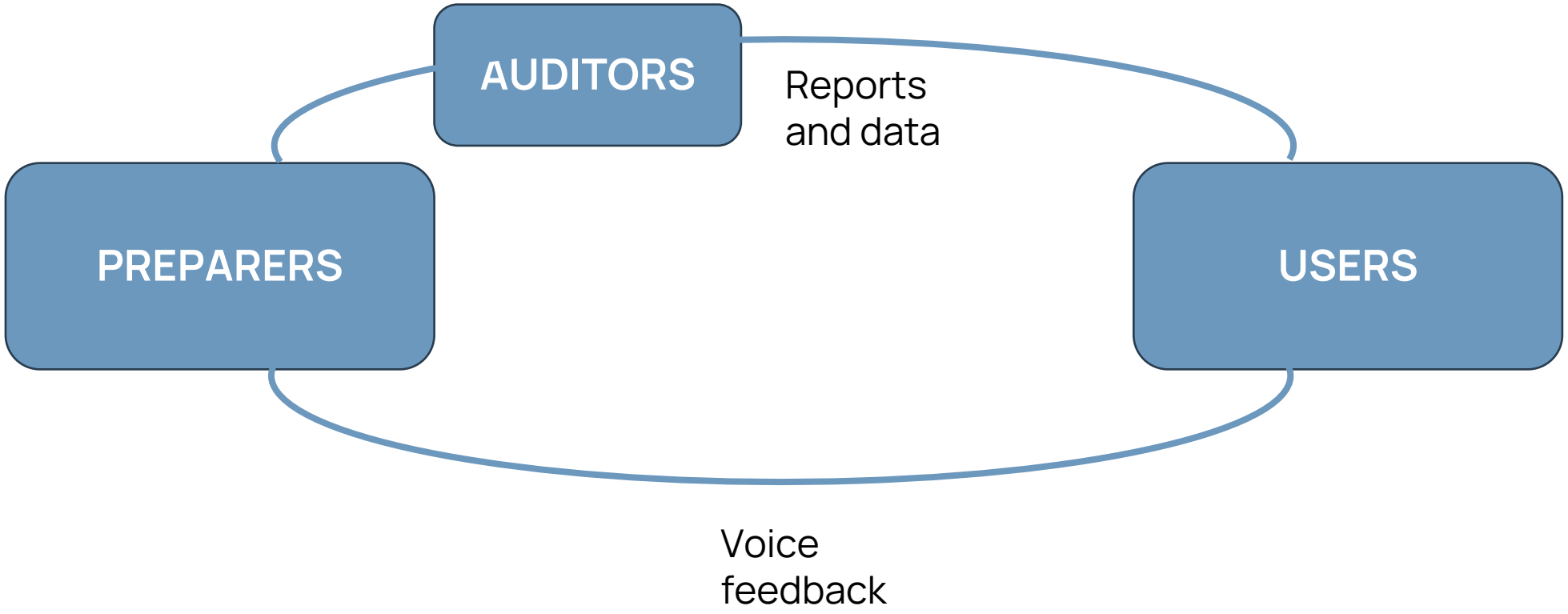
03

AI Implications for accountability and public sector financial reporting

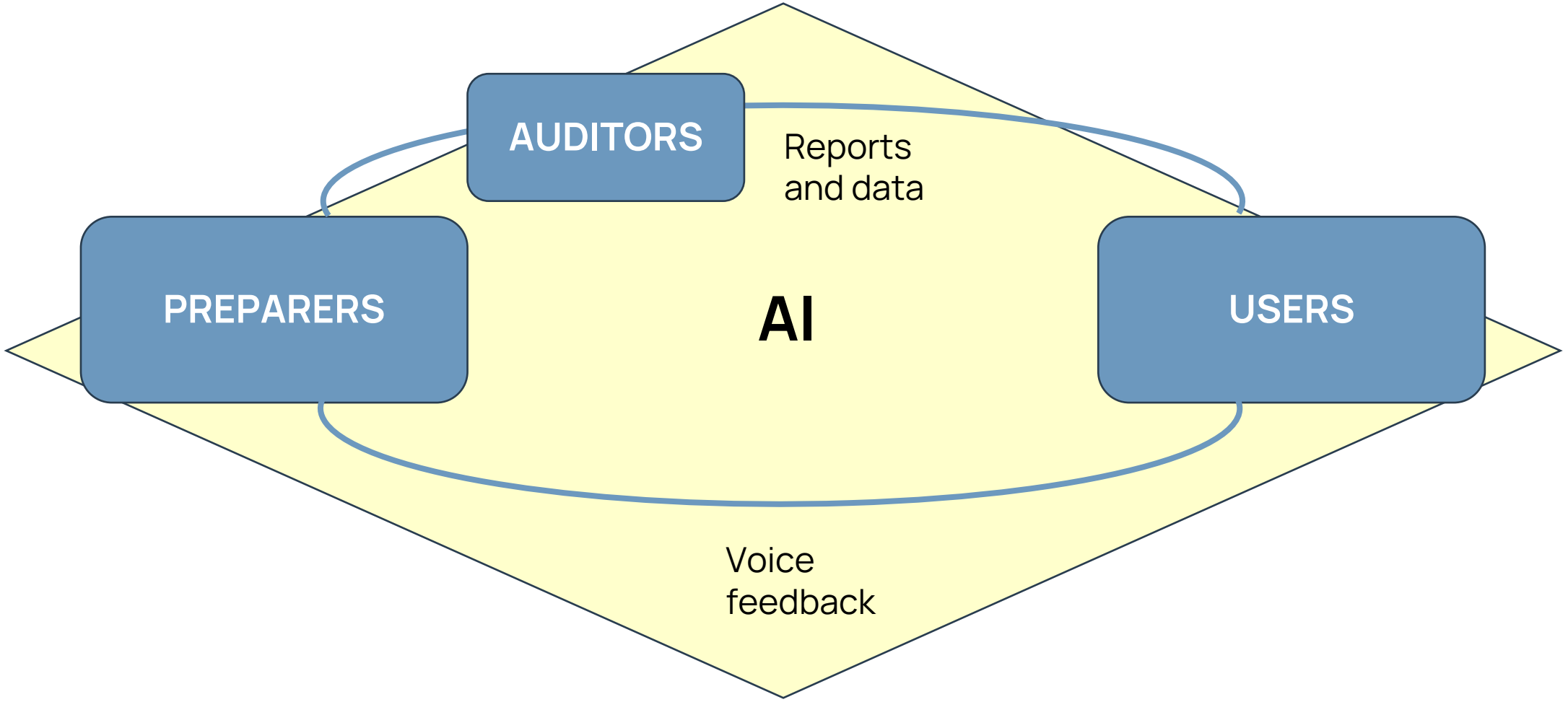
04

Key Questions for IPSASB

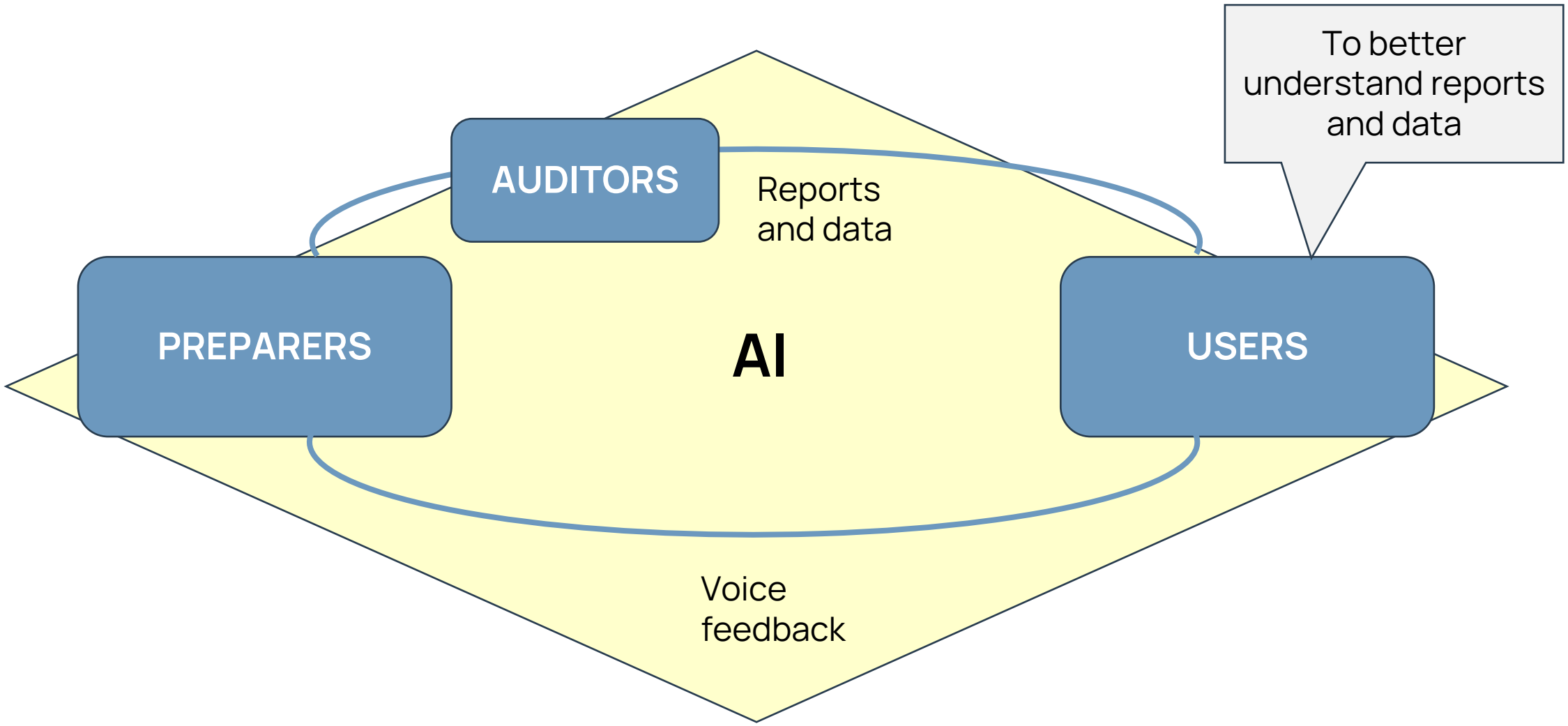
The «traditional» accountability relationship



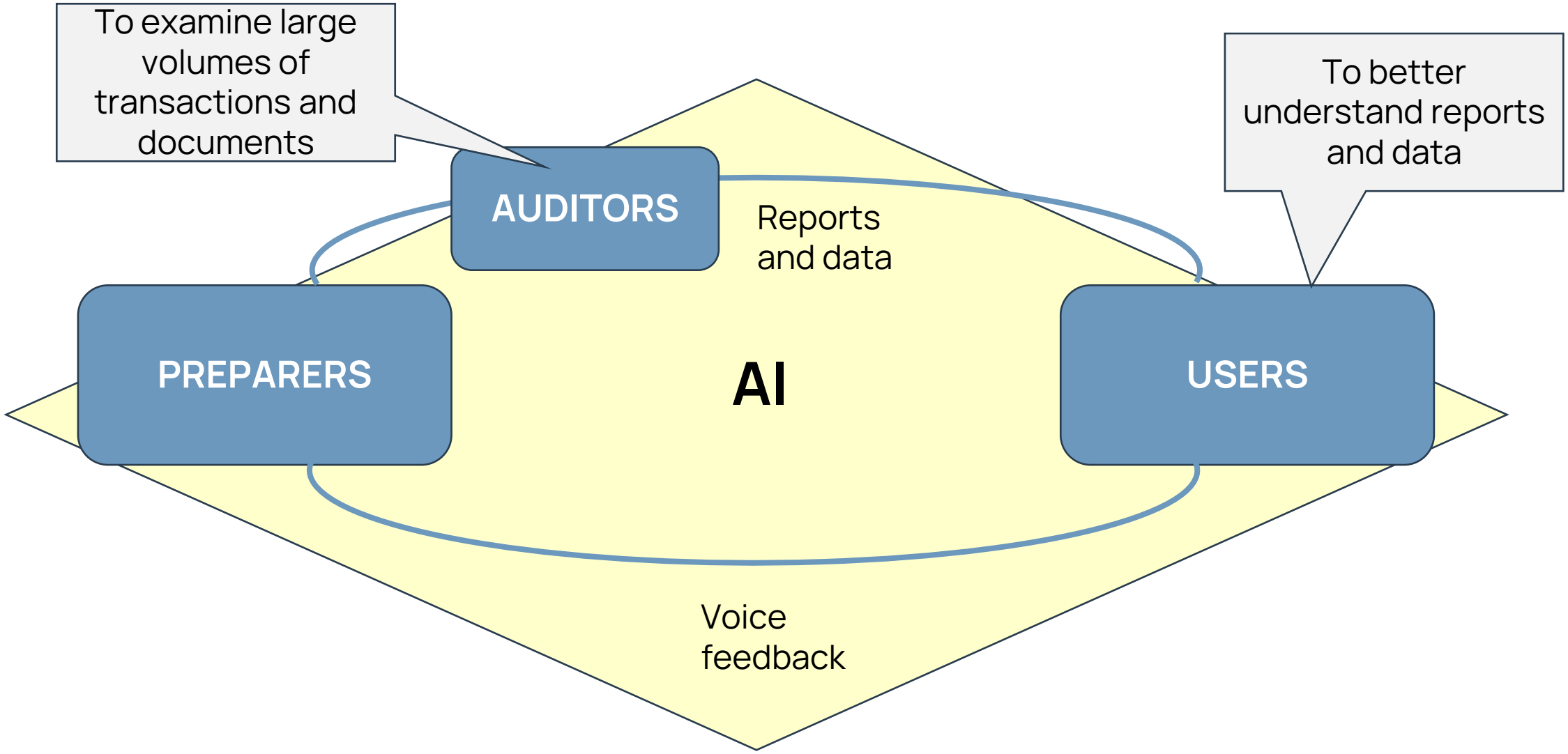
What happens when AI enters the accountability relationship



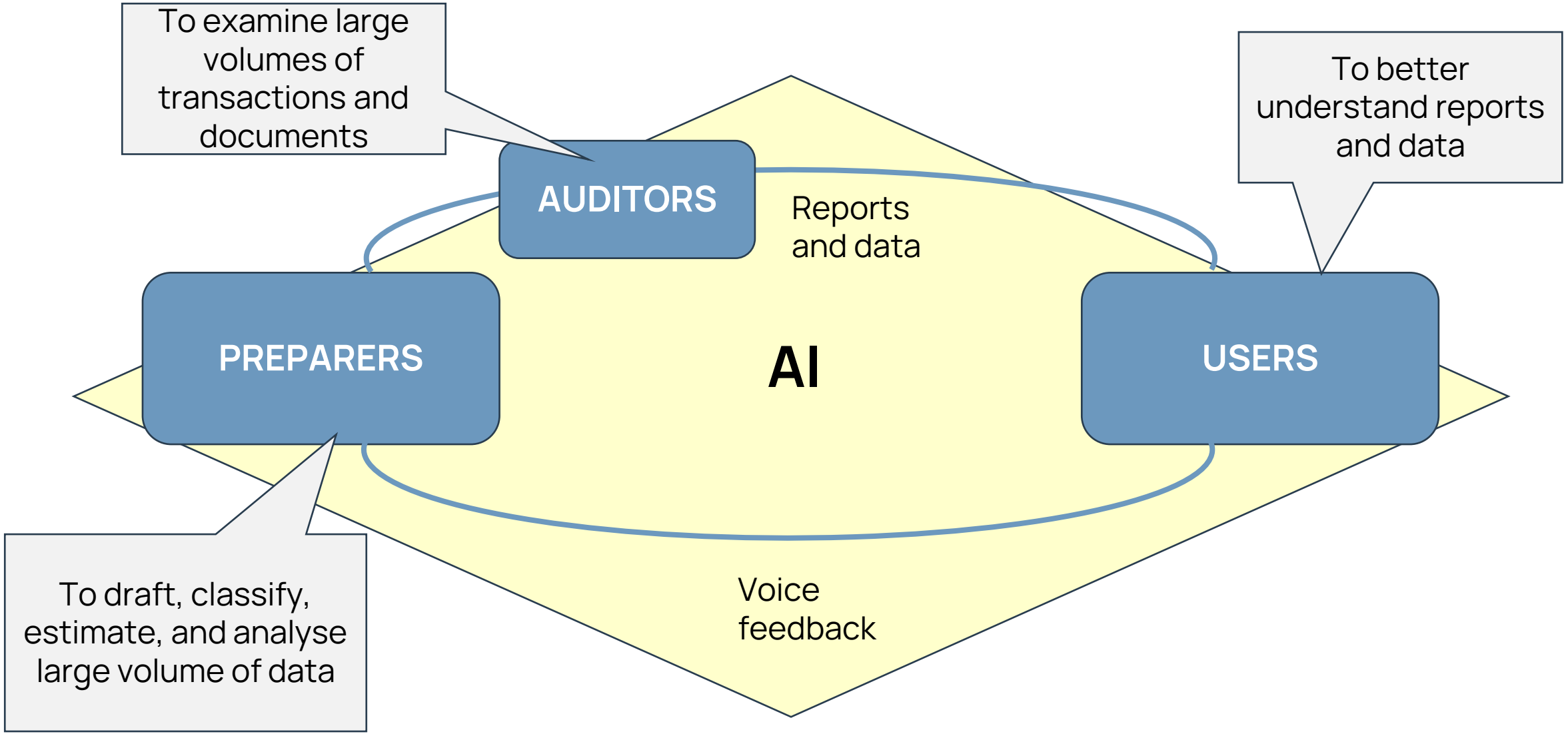
What happens when AI enters the accountability relationship



What happens when AI enters the accountability relationship



What happens when AI enters the accountability relationship



The implications of AI for public sector accountants

Drafting



- Support drafting, summarising, and classification tasks
- **Draft notes** and explanations
- Summarise key distinctions

Estimation



- Help improve forecasting and estimation outputs
- Identify trends and patterns
- Model projections from data

Classification



- Assist in classifying and tagging transactions
- Support coding and classification
- Tag entries in financial statements

Analysis



- Help **spot anomalies** and control a large volume of data
- Detect inconsistencies or variances
- Check calculations and controls

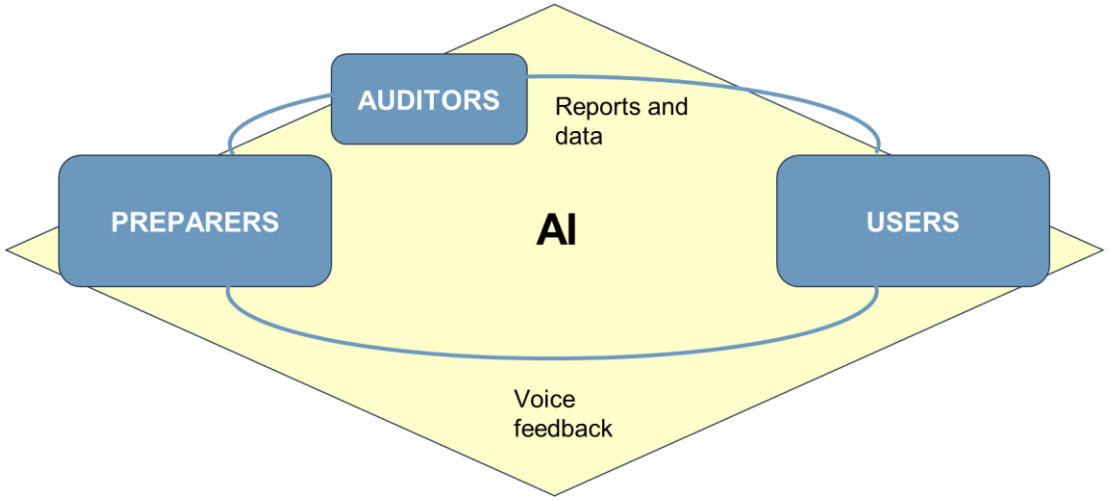
The implications of AI for public sector accountants

- AI can support drafting, summarising, classification, estimation, and data analysis.
- Less time on routine drafting and data handling
- More time on review, interpretation, challenge, and judgement
- But it may also obscure where **judgement** enters the process
- Greater responsibility for explaining how AI-supported outputs were validated

The implications of AI for auditors and oversight bodies

- AI can support anomaly detection, document review, and analytical procedures
- It can improve coverage and speed in high-volume environments.
- But auditors may also need to assess how AI influenced reporting processes.
- Assurance may increasingly require attention to governance, traceability, and human oversight.

The key challenges of an «AI-mediated accountability relationship»



- Process transparency vs output transparency
- Plausible fluency vs verified reality
- Cognitive delegation vs responsibility

Agenda

01

Introduction to the AI

02

Deeper dive into AI models

03

AI implications for accountability and accounting

04

Key Questions for IPSASB

Key risks for accountants in AI-supported environments

- Hallucinations: persuasive but incorrect content
- Bias: distorted outputs based on data or model design
- Over-reliance: reduced professional scepticism
- Weak traceability: difficulty reconstructing how an output was produced
- Responsibility diffusion: unclear boundaries between tool support and human judgement

Conditions for a responsible use of AI in public sector accounting

- Define which uses are permitted, restricted, or prohibited.
- Keep humans accountable for material judgements and external outputs.
- Document prompts, sources, assumptions, checks, approvals, and overrides.
- Apply safeguards for confidentiality, security, procurement, and data protection.
- Build capability through training, testing, and monitoring.
- Use authoritative data sources wherever numbers matter

Issues for IPSASB to Consider

- Where could AI materially affect public sector reporting?
- Which uses of AI are merely operational, and which affect judgement, disclosure or assurance?
- What documentation should exist when AI supports reporting outputs?
- How can **transparency** be preserved when outputs are easy to read but processes are hard to explain?
- What does accountability mean when AI mediates preparation, interpretation, and scrutiny?



POLITECNICO MILANO 1863 | DIPARTIMENTO DI INGEGNERIA GESTIONALE

Thank you!

Deborah Agostino
Deborah.agostino@polimi.it