AZURE OPENAI COST ESTIMATION SERVICE

GENAI IS POISED TO DISRUPT IT BUDGETS IN 2024

Companies are struggling to understand the cost implications of their GenAI initiatives, leading to uninformed GenAI model selection, inadequate business case inputs, and budget overruns

40%
Cloud AI CAGR
2023 › 2030[1]

~13x
Growth in utilisation of
SaaS LLM APIs (ChatGPT)[2]

3.4x
Anticipated growth in 2023
GenAI budgets between
Q3 & Q4 '23[3]

Macro-economic pressures and the promise of productivity and efficiency gains are supercharging the growth of the global cloud AI sector and bloating GenAI budgets demands
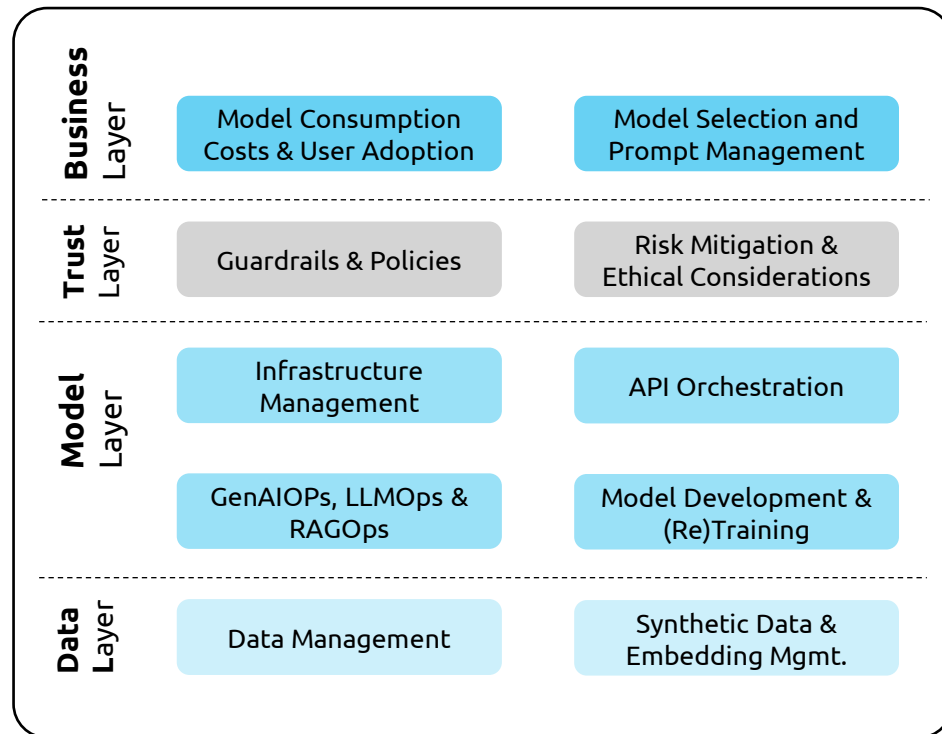
1. Grand View Research global cloud AI market report 2023
2. Growth between end of Nov 22 and May 2023 - Databricks, 2023 State of Data & AI Report
3. IBM, 'Tech spend: How will you pay for it?' Report 2023

COST TRANSPARENCY IS KEY TO UNLOCK VALUE BASED TRADE-OFFS

Companies require better cost estimation models to allow for comprehensive value trade-off decisions across their AI portfolios

GenAI initiaitives have several cost factors which should be considered for every investment decison

Business Layer	Model Consumption Costs & User Adoption	Model Selection and Prompt Management
Trust Layer	Guardrails & Policies	Risk Mitigation & Ethical Considerations
Model Layer	Infrastructure Management	API Orchestration
	GenAIOPs, LLMOps & RAGOps	Model Development & (Re)Training
Data Layer	Data Management	Synthetic Data & Embedding Mgmt.

Increased awareness and application of cost estimation models can greatly reduce budget vairance and improve value realisation

Effective value-based selection of GenAI resources

More accurate forecasting & budgeting

Improved model selection & investment decisions

CAPGEMINI AZURE OPENAI COST ESTIMATION SERVICE PROVIDES CLARITY

Using our propriety Azure OpenAI cost estimation tool and our consulting advisory experience we can help you select, size and compare your OpenAI services

For Language and Fine-tuning models, select from a list of example scenarios to help you size the likely inference usage from your users

Choose from a selection of Model types, including Language, Fine-tuning, and Embedding

Based on likely usage inputs for your chosen model and user base, the tool will generate 5-year cost estimates. You can also compare costs for the same usage across other available OpenAI models to support decision making and export the estimates to CSV for further analysis
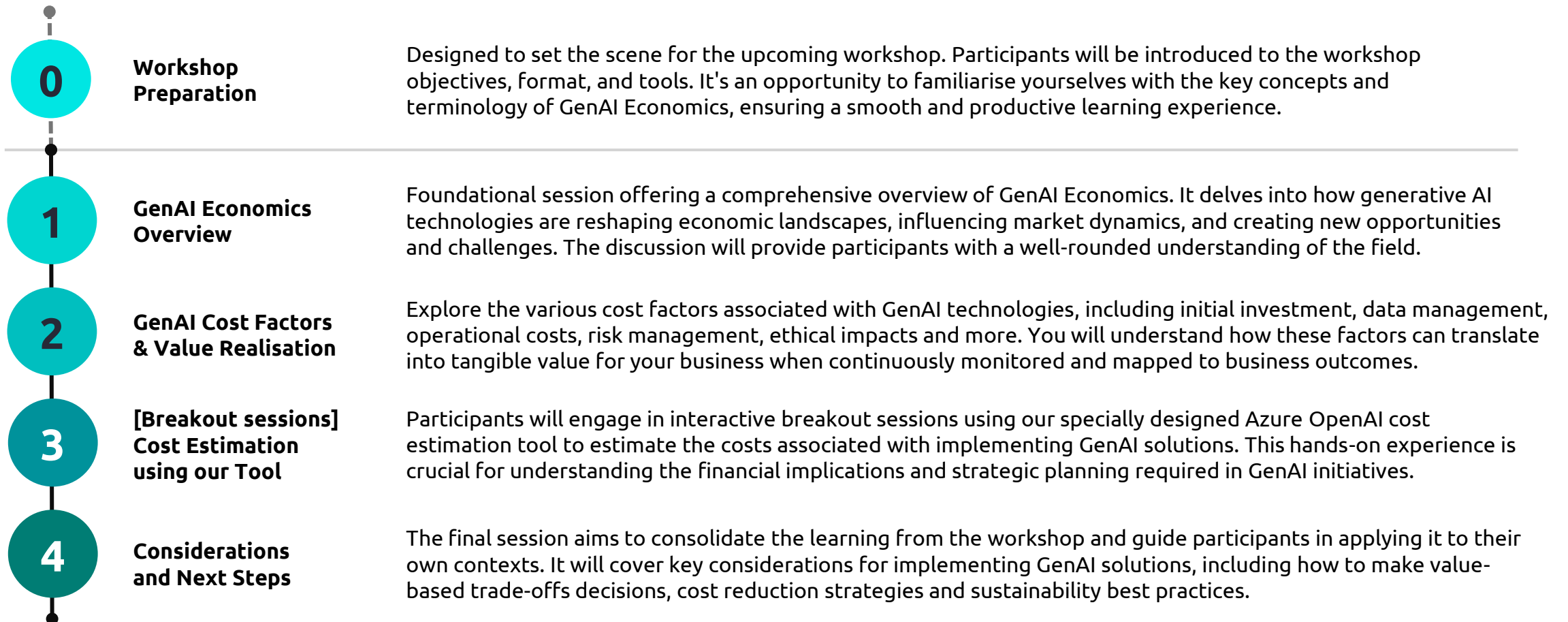
'24 Feature Roadmap

- Multi-model cost tallying functionality

- Expansion of training and help documentation

- Inclusion of self-hosted cost estimation module

- Added functionality to include cost factors such as data preparation, operational considerations, etc

- And more..

HOW WE ENGAGE WITH YOU

Tailored half day GenAI economics workshop designed to support you understand the cost factors involved in the selection, sizing and comparison of Azure OpenAI services

0 — **Workshop Preparation**

Designed to set the scene for the upcoming workshop. Participants will be introduced to the workshop objectives, format, and tools. It's an opportunity to familiarise yourselves with the key concepts and terminology of GenAI Economics, ensuring a smooth and productive learning experience.

1 — **GenAI Economics Overview**

Foundational session offering a comprehensive overview of GenAI Economics. It delves into how generative AI technologies are reshaping economic landscapes, influencing market dynamics, and creating new opportunities and challenges. The discussion will provide participants with a well-rounded understanding of the field.

2 — **GenAI Cost Factors & Value Realisation**

Explore the various cost factors associated with GenAI technologies, including initial investment, data management, operational costs, risk management, ethical impacts and more. You will understand how these factors can translate into tangible value for your business when continuously monitored and mapped to business outcomes.

3 — **[Breakout sessions] Cost Estimation using our Tool**

Participants will engage in interactive breakout sessions using our specially designed Azure OpenAI cost estimation tool to estimate the costs associated with implementing GenAI solutions. This hands-on experience is crucial for understanding the financial implications and strategic planning required in GenAI initiatives.

4 — **Considerations and Next Steps**

The final session aims to consolidate the learning from the workshop and guide participants in applying it to their own contexts. It will cover key considerations for implementing GenAI solutions, including how to make value-based trade-offs decisions, cost reduction strategies and sustainability best practices.

WHAT MAKES US UNIQUE?

20+
GenAI Use Cases implemented

5+
Generative AI-based solutions in place

8+
Partners to deliver Generative AI solutions

Capgemini
Dedicated **AI Futures Lab**

20 000+
Experts across domains & industry sectors

30 000+
Data & AI Consultants

OUR CROSS-INDUSTRY FOOTPRINT

Automotive

Aerospace & Defense

Banking

High Tech

Life Sciences

Telecomms

Energy & Utilities

Insurance

Public Services

Consumer Products & Retail

A PROVEN DATA & AI PORTFOLIO EXPERTISE

- **Multiple years of experience** in deploying scaled Gen.AI solutions in production
- **Intellectual Property and accelerators** built-in house or in collaboration with key partners
- **Research-based Technical Thought Leadership** incl. Prompt Hydration and Privacy Preserving AI techniques
- **Proven-record in data platforms** with own IP (IDEA by Capgemini) which position us as a preferred partner for AI implementation at scale
- **Ability to guide through ethical and legal** compliance complexity such as AI Act with Capgemini Invent

Dedicated Gen.AI partnership with: **Microsoft Azure**