AI ANALYTICS & AUTOMATION

Episode 1: AI for Executives

Nick White

May 2024

Agenda

1/ The Reality
2/ Talk the Talk
3/ Think Big, Start Small

Objectives of this Session

As an executive you feel empowered to drive innovation with AI.

As a data professional you are equipped to educate your organization about AI.

As a user you understand how to use AI better.

1/ The Reality

About AI

AI for productivity…now

14%

Productivity gains realized by customer service agents

Source: NBER

40%

Performance improvement with highly-skilled non-technical workers

Source: MIT Sloan

58%

5-year productivity gains for developers

Source: IBM Study

ORIGIN®

AI will transform…soon

- Enterprise applications of GenAI are not simple but are highly accelerated by the availability of powerful foundational models
  - Since 2015, OpenAI has secured more than $12 billion in investments, with GPT-4 Turbo costing $63 million in computing alone

- Organizations must invest in productivity and transformation with GenAI right now to win in their space
  - Making the right investments is complicated as the secondary GenAI market is grows


All in!

AI is not easy

85%

AI projects never make it to production.

Source: Gartner

96%

AI projects run into data quality, labeling and building model confidence.

Source: Dimensional Research

81%

Training AI with data is more difficult than expected.

Source: Analytics India

ORIGIN®

Risks presented by AI

The unique risks presented by AI have both commercial and ethical dimensions.

Bias

AI is trained on large amounts of data, which can contain biases and perpetuate existing societal inequalities and prevent innovative thinking.

Fraud

AI can be used to generate convincing fake text or media, which can be used to spread misinformation or open organizations up to internal and external fraud.

Breaches

AI requires huge volumes of data to be effective potential exposing more private data subject to security breaches.

Consumption

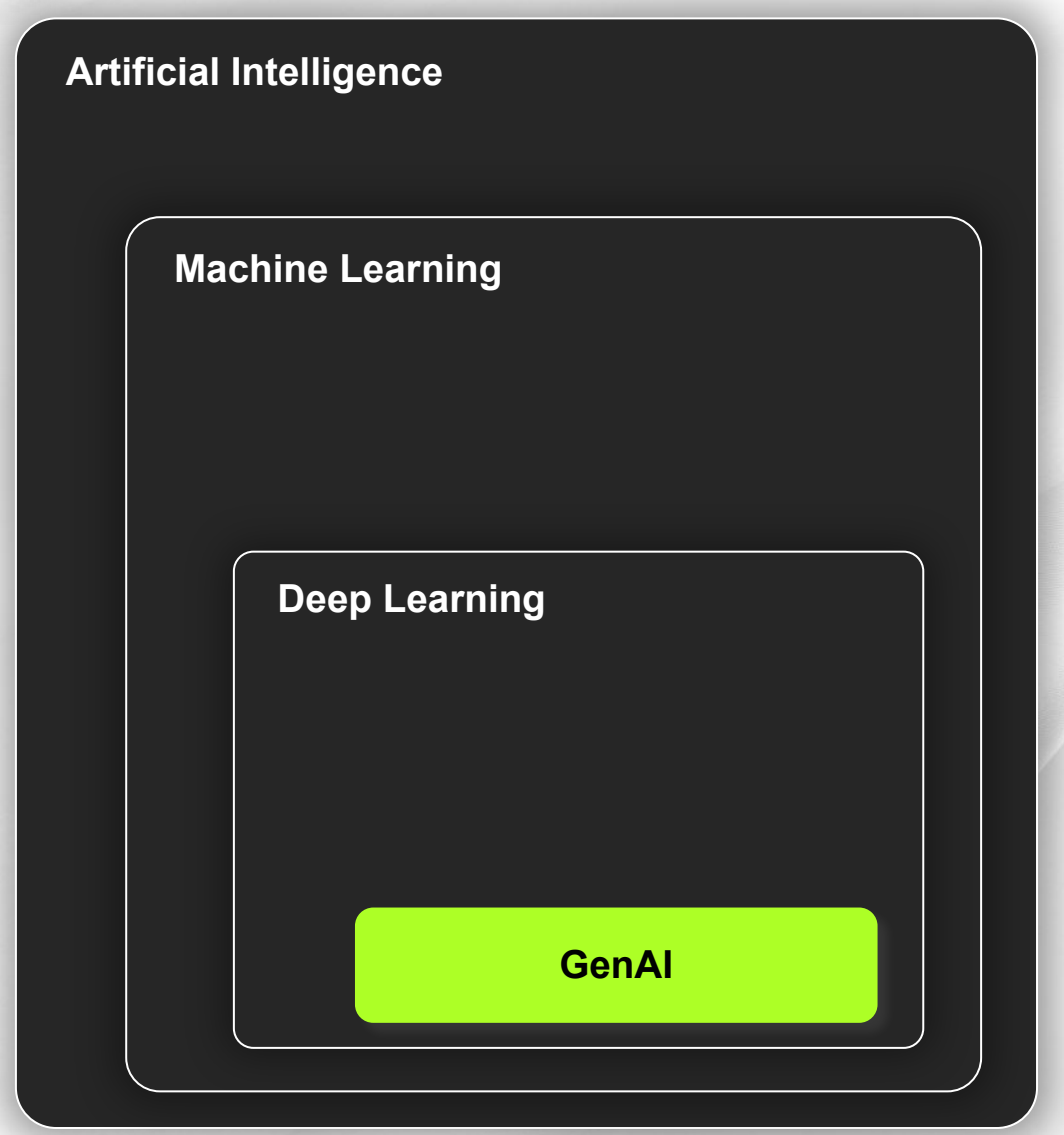
AI is a complex and resource-intensive technology that requires significant computing energy consumption and costs if not managed efficiently.

Hallucinations

While AI can generate highly convincing text, images and audio, it is not always accurate or reliable opening organizations up to costly errors.


Stay away!

AI is old



1956

Artificial Intelligence
The field of computer science that seeks to create intelligent machines that can replicate or exceed human intelligence

1997

Machine Learning
Subset of AI that enables machines to learn from existing data and improve upon that data to make decisions or predictions

2017

Deep Learning
A machine learning technique in which layers of neural networks are used to process data and make decisions

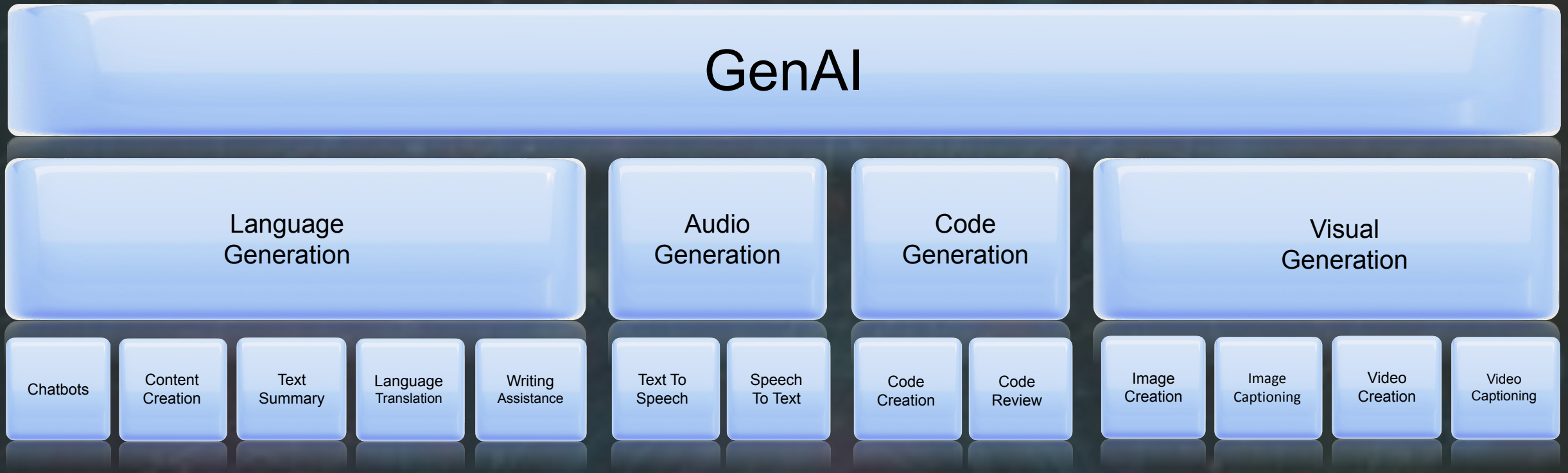
2021

Generative AI
Create new written, visual, and auditory content given prompts or existing data

ORIGIN

GenAI Capabilities

Create various types of content including language, visuals, audio and code using AI.



ORIGIN

The Deep Learning Models Behind GenAI

Language Model
A deep-learning AI model designed to understand and predict text using massive amounts of parameters and training data.

Examples: GPT, PaLM, LLaMA

Computer Vision
A deep-learning AI model designed to process and interpret visual data, typically images or videos.

Examples: GPT-Vision, ACS Computer Vision, Google Cloud Vision

Automatic Speech Recognition (ASR)
A deep-learning AI model designed transcribe spoken language into written text. May be called Speech-to-Text.

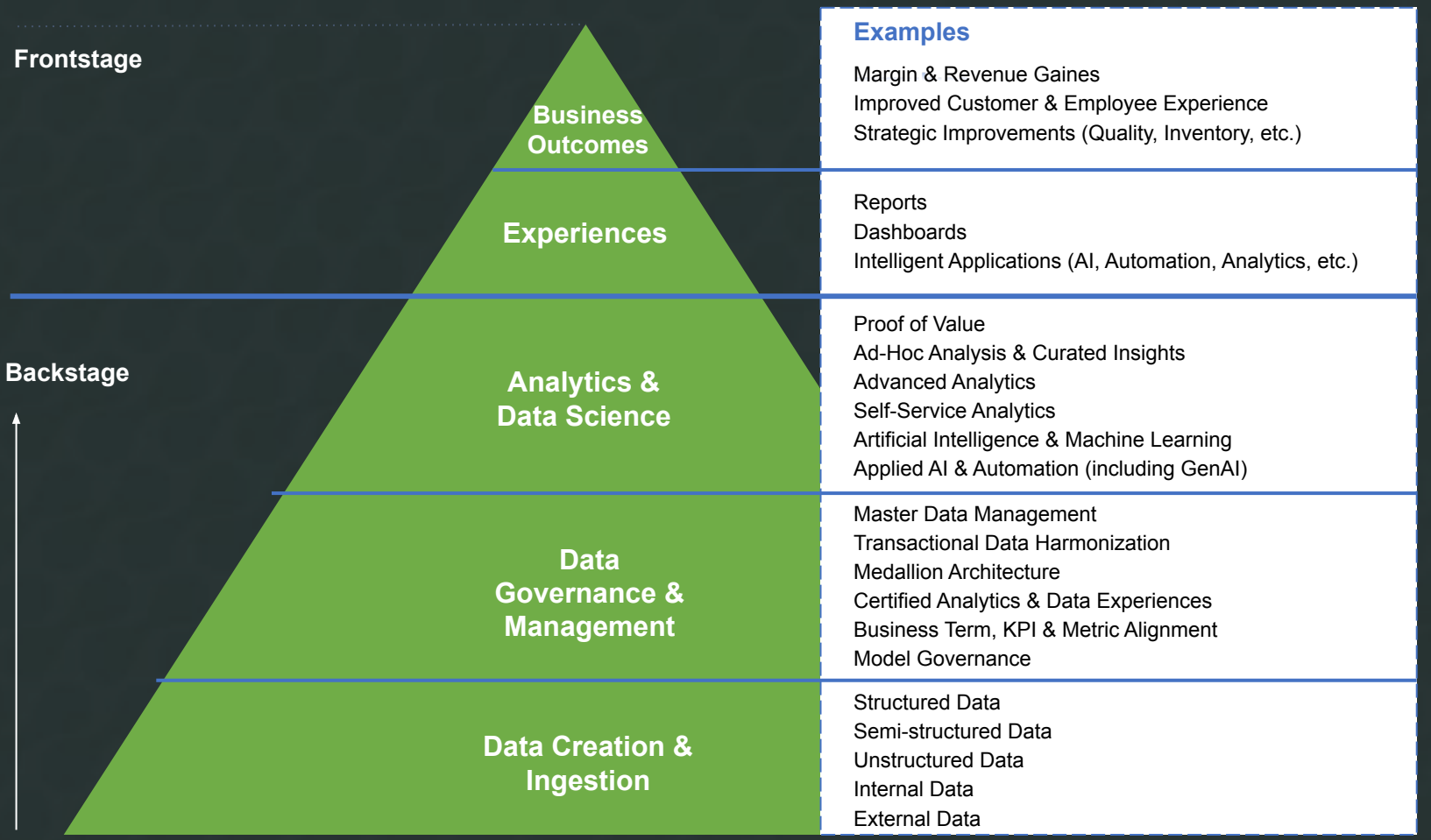
Examples: Whisper, Azure Speech Service, Google Speech-to-Text

Multimodal Model
A deep-learning AI model that can process and understand multiple types of data modalities such as text, video, images, audio and TBD others.

Examples: GPT-4o, Gemini

ORIGIN®

AI is one aspect of the data-driven organization



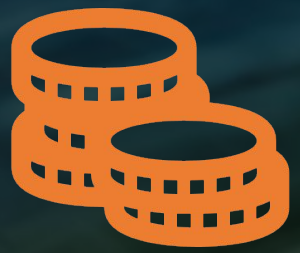
ORIGIN


Let’s try.

AI Value Realization

Grow Faster

At constant cost, increase ability to generate and deliver more revenue



Get Leaner

At constant revenue, reduce costs to capture value of gained productivity

Innovation Curve for AI

5%

Leading Edge Organizations

- Less than 5% of total organizations
- Typically have internal research teams and bandwidth to experiment with new technologies to determine opportunities
- Often partner with their technology partners and management consultancies/agencies in “bleeding edge” work

35%

Proactive Organizations

- About a third of companies keep an open eye towards potential technology opportunity and disruption.
- They don’t often do their own basic research but as soon as some basic research is available, they engage with consultancies and their partners to determine fit and opportunity

60%

Reactive Organizations

- The remainder of organization take a reactive perspective towards technology disruption.
- This can be an intentional position to wait for others to invest in research and experimentation. This behavior can also originate from a lack of overall technology maturity, which may or may not be intentional.

Remember, GenAI is just a tool

Although GenAI models can accelerate time and cost to value with AI, the amount of value is subject to the same factors as other AI applications.

Strategy	**AI initiatives must be aligned to specific business goals**
Expertise	**Practitioners and end-users must have the skills to build and use AI**
Data Quality	**Garbage in, garbage out applies with AI, even GenAI**

ORIGIN

2/ Talking AI

Understanding the Lingo

AI is Everywhere (and has been for a while)

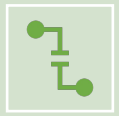
Service Level	Description	Intended User	Customization	Value Realization
AI Applications	A standalone application to interact with an AI model for out of the box functionality (e.g. ChatGPT, Co-Pilot Web, etc.).	End-User	None	Up-skilling end-users to leverage AI experience with managed and curated data, content and prompts.
Applications With AI	AI is embedded in other applications to improve the user experience with greater utility, ease of use and efficiency (e.g. Office 365, Gmail, etc.).	End-User	None	
No-Code AI Platforms	Business users can create their own AI applications with an easy to use, no-code platform (e.g. Co-Pilot Studio, OpenAI GPTs, etc.).	Power End-User & AI Solution Developers	Low	**All the above** and light tuning & integrations.
AI Applied Models	Pre-tuned AI models exist to solve for the most common use-cases (Search, Document OCR, Chatbots, etc.). Still require customization by developers.	AI Solution Developers	Medium	**All the above** plus create integrations and tuning with managed content and data available in a UI.
New AI Foundation Models	Pre-trained models can be leveraged for unique applications of GenAI or unstructured data mining (e.g. GPT, LaMDA, PaLM, BLOOM, etc.).	AI Solution Developers & Data Scientists	High	
AI Platforms	Provide a comprehensive environment for building and operationalizing AI solutions (e.g. Databricks, Microsoft Fabric, Snowflake, etc.).	Data Scientist, Analysts & Engineers	Highest	**All the above** and create and train custom models.

ORIGIN®

How AI Foundation Models are Created

Unsupervised Pre-Training with sources such as books, websites, and articles.

Reinforcement Learning using AI and humans to provide feedback to improve decision-making.



Adversarial Testing creating challenging scenarios or inputs to test the model's behavior under stress.

General Availability via a natural language or graphical interfaces and/or APIs.

ORIGIN™

Foundation AI Model Definitions

Pre-Training
Process in machine learning and natural language processing where a model is initially trained on a large dataset before being fine-tuned for specific tasks. Unsupervised: no human labeling; supervised: human labeled data.

Fine Tuning
The process of adjusting the parameters of a pre-trained model to adapt it for a specific task or domain. The more specific a task, the more fine-tuning in necessary.

Reinforcement Learning (RL)
A machine learning technique that trains software to make decisions to achieve the most optimal results through reward and punishment. This can be used to make a GenAI application better over time.

Adversarial Testing
A method used to systematically evaluate a machine learning model by intentionally providing it with malicious or inadvertently harmful input.

Prompting
Instructions to steer a foundational model towards generating specific outputs without the need for extensive retraining. This can be done before end-user interaction and is called prompt engineering or by the end-user.

Grounding Data
Information that is use-case specific, relevant, and not available as part of the foundational GenAI model's trained knowledge. Used in Retrieval-Augment-Generation architecture.

ORIGIN

Key Considerations AI Foundation Models

Availability
The availability of a foundation model refers to how accessible the model is to the public or specific organizations. It can be classified based on whether the model is open source or closed source.

Size
The size of a foundation model refers to the number of parameters it contains. Parameters are the elements of the model that are learned from training data and are used to make predictions. Models can range from millions to hundreds of billions of parameters.

ORIGIN®

AI Foundation Model Availability

Open Source

Definition:
Open-source foundation models are AI-trained models whose architecture and code are publicly available. They allow researchers and developers to use, modify, and distribute the models freely, fostering collaboration and innovation.

Use Cases:
- Research and Development: Open-source models are extensively used in academic research, enabling the scientific community to validate and extend existing models.
- Customization: Organizations can customize these models to meet specific needs without licensing restrictions.
- Transparency and Trust: Open-source models promote transparency, allowing users to understand how the models work and ensuring trust in AI applications.

Example Models:
- LLaMA (Large Language Model Meta AI): Developed by Meta, a family of large language models designed to be efficient and accessible for research and development.
- BLOOM (BigScience Large Open-science Open-access Multilingual Language Model): A multilingual language model developed collaboratively by the AI community.
- GPT-2: An earlier version of the GPT-3 model, released by OpenAI, known for its capability in text generation and completion tasks.

Closed Source

Definition:
Closed-source foundation models are proprietary AI-trained models whose architecture, code, and training data are not publicly available. These models are typically developed by companies and are accessible through paid licenses or subscription services.

Use Cases:
- Commercial Applications: Businesses use closed-source models for their commercial products and services, leveraging the advanced capabilities provided by the proprietary models.
- Data Privacy and Security: Companies may prefer closed-source models for sensitive applications where data privacy and security are paramount.
- Exclusive Features: Closed-source models often come with exclusive features and optimizations that are not available in open-source alternatives.

Example Models:
- GPT-4: The latest version of the popular foundation model, excelling in text generation, translation, and various other NLP tasks, available through OpenAI or Microsoft Azure.
- Gemini (formerly PaLM): A family of multimodal large language models developed by Google DeepMind, serving as the successor to LaMDA and PaLM 2.
- Amazon's Alexa: Amazon's voice assistant technology, which is built on proprietary foundation models, providing robust speech recognition and natural language understanding.

ORIGIN

AI Foundation Model Size

Large Foundation Models

Definition:
Large foundation models are massive AI-trained models that utilize substantial amounts of data and computational resources. They serve as the starting point for developing more advanced and complex models.

Use Cases:
- Broad Natural Language Processing (NLP): Large foundation models excel in NLP tasks such as text generation, translation, sentiment analysis, and question answering.
- Computer Vision: These models can also handle image classification, object detection, and other vision-related tasks.

Example Models:
- GPT-4 Turbo: The most powerful version of the popular foundation model, reasoning across audio, vision and text in real time available directly from OpenAI or through Microsoft Azure.
- Gemini Ultra: A large multimodal model developed by Google to handle a wide range of tasks across audio, images and text.
- Llama-3 70B: A large language model developed by Meta, featuring 70 billion parameters. It's designed to perform a wide range of natural language processing tasks and is known for its impressive performance.

Small Foundation Models

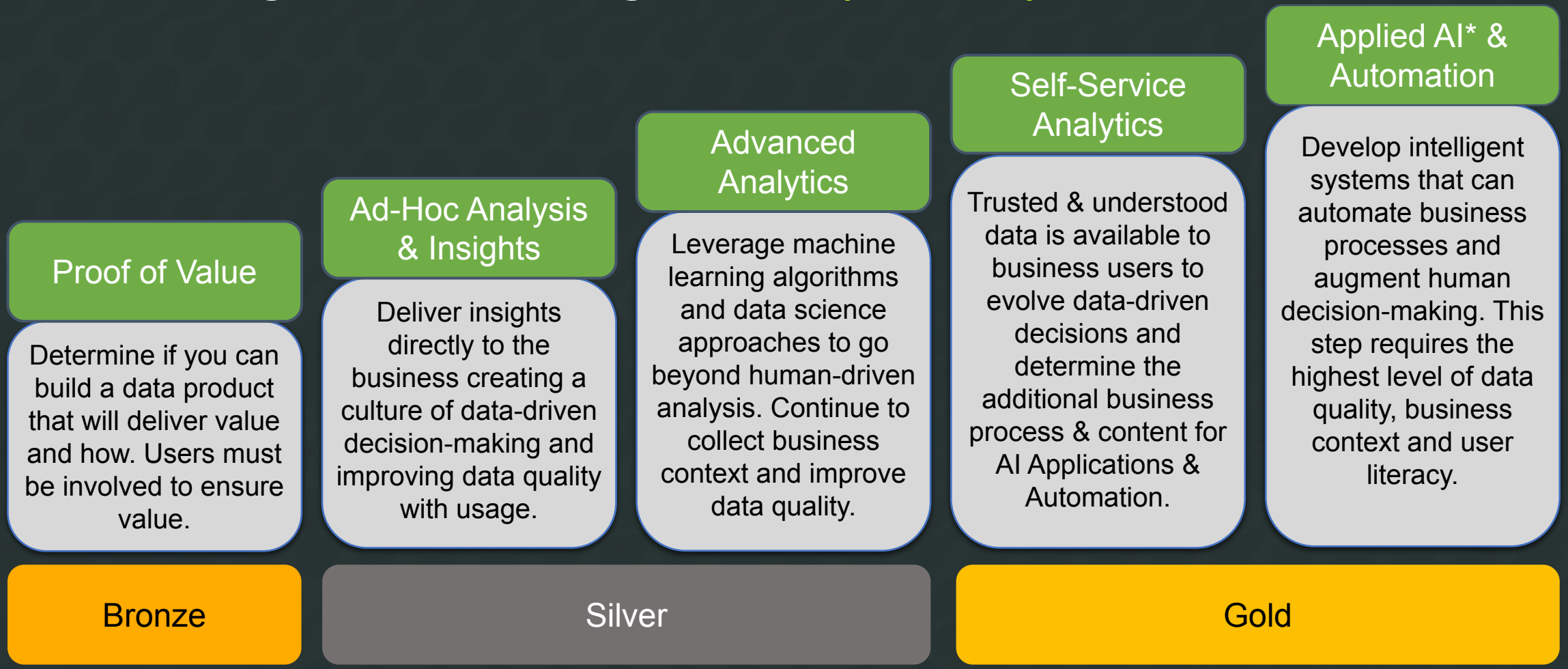
Definition:
Small foundation models (SFM) have fewer parameters (typically 12 billion or less) compared to their larger counterparts. They are resource-efficient (cheaper!) and easier to manage.

Use Cases:
- Chatbots and Virtual Assistants: SFMs are suitable for basic interactions and concise responses.
- Resource-Constrained Environments: SFMs work well in scenarios with limited computational resources (edge-computing - mobile).
- Domain Intensive Applications: If your use case requires a ton of specific context, fine-tuning an SFM may be the right choice.

Example Models:
- Phi-2, Phi-3 & Orca-2: Microsoft's small language models.
- Mistral's 7B: Designed for enterprise applications by Mistral.
- Gemini Nano: Google's compact model for efficiency.
- Llama-3 8B: Small language model from Meta.
- Dolly 2.0: Databricks' small foundational model.

ORIGIN

Garbage in, Garbage Out (GIGO)



* Includes Generative AI

ORIGIN

3/ Get Started

Think Big, Start Small

Create a culture of innovation

Cross-functional Teams
A culture of innovation promotes collaboration, enabling cross-functional teams to work together and drive the success of AI projects

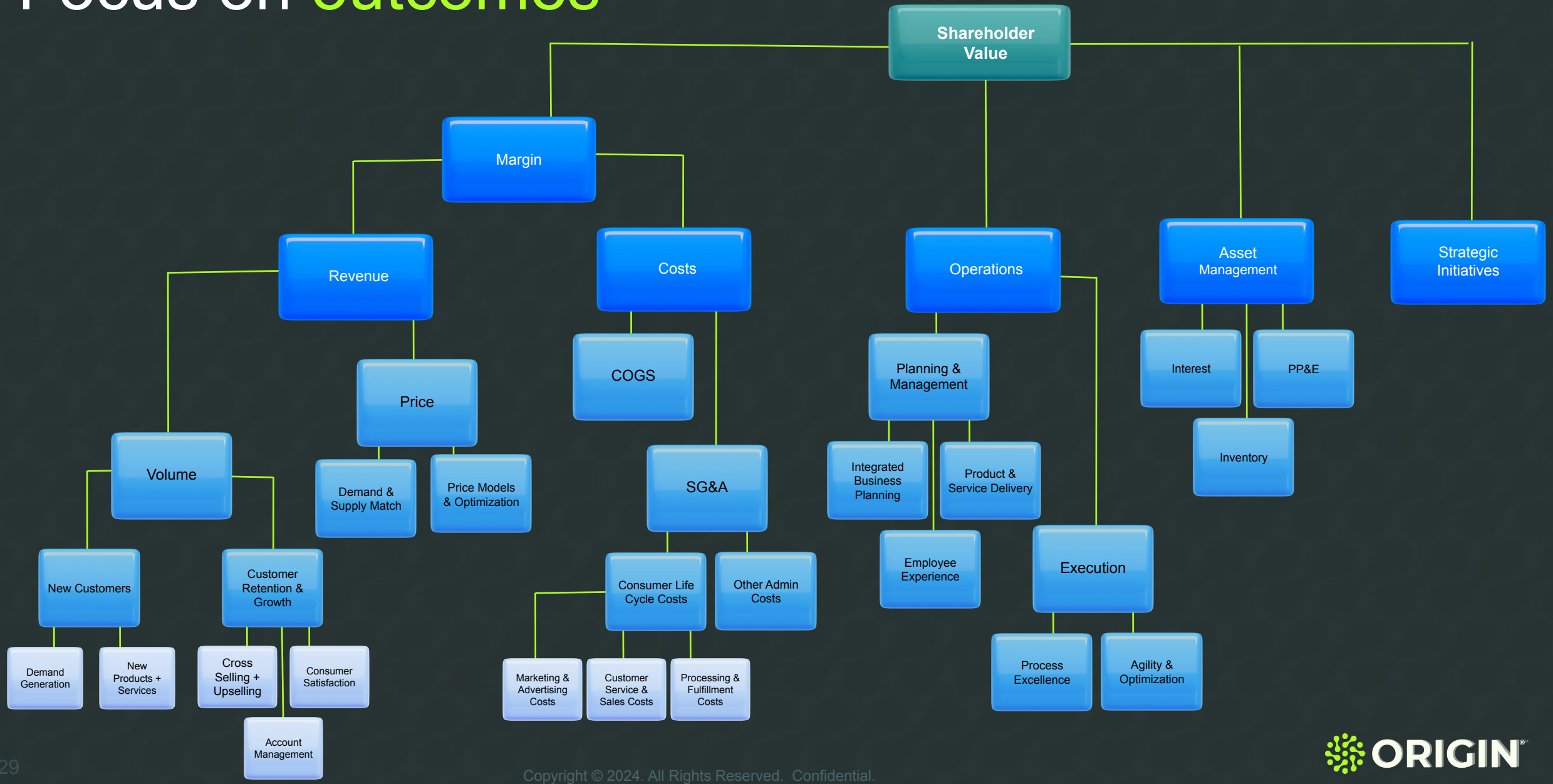
Continuous Learning
Innovation culture fosters an environment of continuous learning and adaptation, allowing AI projects to constantly evolve and improve

Design-Thinking

Innovation culture promotes a service design mindset, allowing AI projects to quickly respond and adapt to changing conditions

ORIGIN

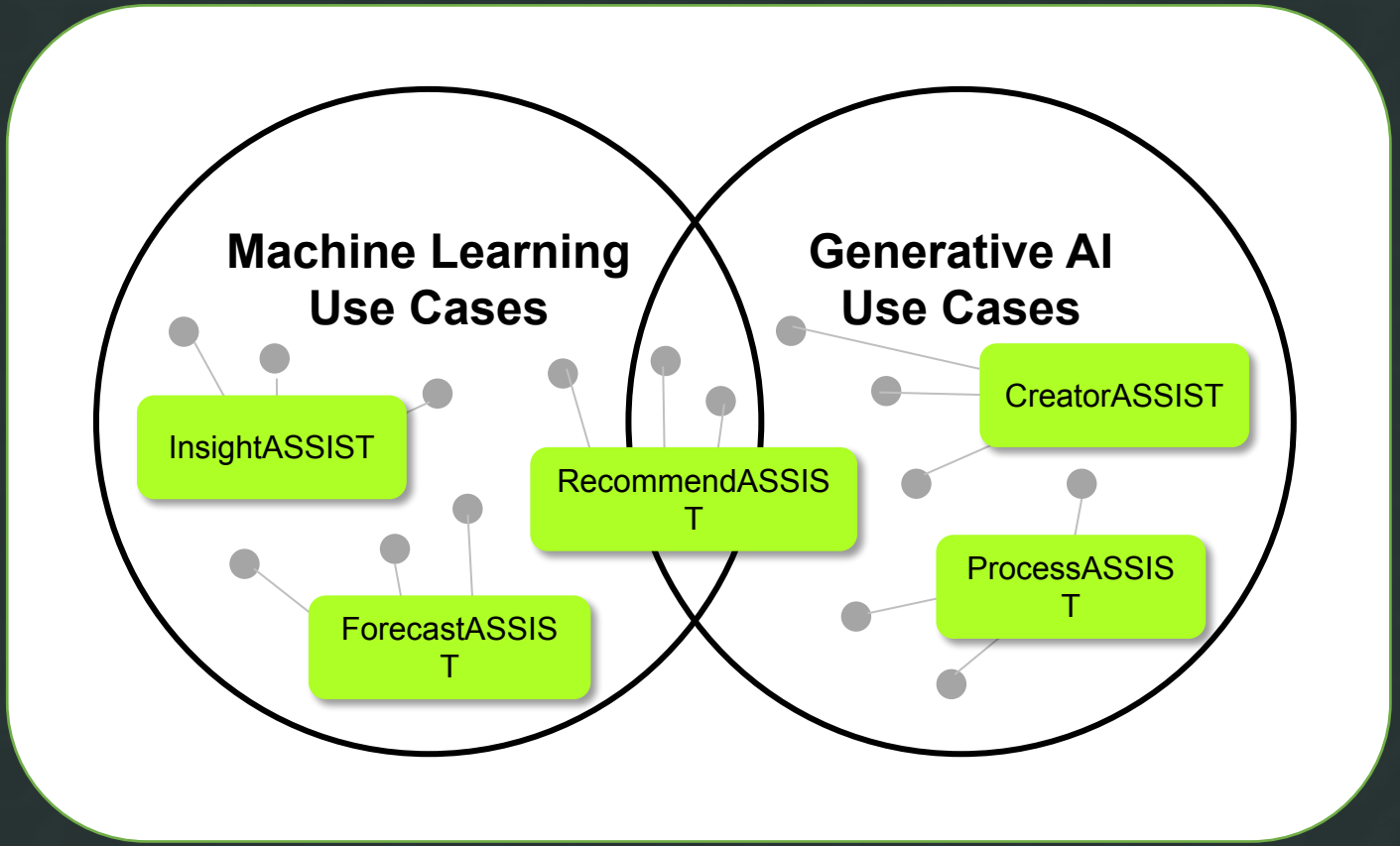
Focus on outcomes



ORIGIN®

Use a Holistic Approach

Determine how AI may help drive outcomes at a broader level.



InsightASSIST
Where can you make it easier to find important insights to drive better decisions with AI?

ForecastASSIST
How could ML provide better guesses to increase efficiencies and revenue?

RecommendASSIST
Are there complicated decisions that AI could help make easier?

CreatorASSIST
Do you need to create more content in a cheaper manner?

ProcessASSIST
What workflows could benefit from an AI agent providing assistance?

Start with an AI Strategy

Implement a AI Strategy one use-case at a time with an agile mindset.

Prioritize Use-Cases

Identify the specific use-cases where AI is the appropriate tool to add value in line with business objectives. Prioritize use-cases by value, effort urgency and competitive differentiation.

Use-Case Analysis

Identify the correct AI services, additional integrations and data required to deliver AI solutions that deliver iterative value against use-cases. Perform a buy vs. build analysis as part of this strategy.

Assess Expertise

Determine the expertise needed to both build and use a AI solution. Consider training, change management and bringing in outside experts to help close gaps.

Create Roadmap

Define how to deliver value with features on the user interface, prompt tuning, improved grounding data, adding integrations and growing organizational data & AI maturity in an agile manner over time.

Deliver Collaboratively

Business, technology and data should work in concert to deliver valuable, scalable and trustable AI experiences for end users. A consistent captive team is ideal.

ORIGIN®

The Agile Approach to AI

Think Big (Envision)

Vision
Create a customer service chatbot that helps customers get answers to questions quickly with the least amount of friction.

Users
- Customer service representatives
- Members

Outcomes
- Reduce number of customer service calls
- Reduce handle time of customer service calls
- Increase customer satisfaction

Features
- Chatbot interface
- Integration with internal app(s)
- Integration with external app(s)
- Accurate answers or no answers
- References for answers
- Feedback on answers from users
- Two-way chat
- Authentication & security

Data & Content
- Internal knowledge base & FAQs
- Appropriate external knowledge base
- Customer interaction history & logs
- Customer account information
- Chatbot performance history

Start Small (MVP)

Users
- ***Veteran*** customer service representatives

Features
- Chatbot interface
- Accurate answers or no answers
- References for answers
- Feedback on answers from users

Data & Content
- Internal knowledge base & FAQs ***for loans only***

ORIGIN™

An AI strategy needs a data strategy

A good data strategy is defined, designed & delivered.

Define
(Strategy)

Envisioning the future and understanding the present to grow data maturity while delivering iterative & scalable value.

- Identify and prioritize use-cases (Now, Near, Next)
- Target state architecture & organization definition
- Data maturity assessment (People, Process, Tech)
- Gap analysis
- Initial platform & tooling recommendations
- Initial organizational recommendations
- Initial program roadmap

Design
(Tech Foundations)

Validate, design & refine the platforms, tools & patterns used to build valuable data products.

- Data rationalization & migration
- Data processing pattern identification
- Storage & compute optimization
- Data governance & quality tooling
- Data science platforms
- Self-service platforms (dashboarding, analysis)
- Data operations & services

Deliver
(Org Enablement)

Validate, design & refine the processes, frameworks and training needed to build valuable data products.

- Team model activation
- Process & framework definition
- Data stewardship & governance
- Data product management
- Data literacy
- Change management

ORIGIN®

A data strategy must create a source of truth

Desptie the fragmented data landscape.

Front Stage

Behaviors, Preferences & Experiences

Consumer Sales, Merchandising & Inventory

Back Stage

Products, Operations & Commercialization

Capacity, Inventory & Logistics

ORIGIN

Safety first

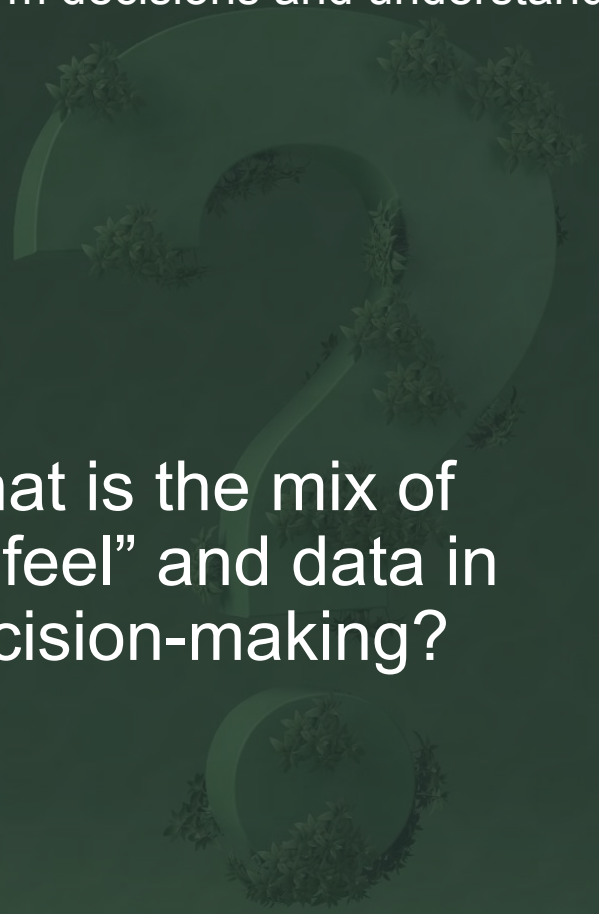
Whether out of the box, bespoke or somewhere in-between, create a safe environment for AI applications.

Documentation and Versioning	Access Control	Testing and Validation	Monitoring and Alerts	Ethical Considerations	Traceability and Transparency
• Maintain detailed documentation for each AI model, including its purpose, features, and assumptions. • Version control helps track changes, making it easier to understand model evolution and ensure reproducibility.	• Define who can access, modify, and deploy AI models. • Implement role-based access controls to restrict unauthorized changes or misuse.	• Regularly test and validate AI models against new data. • Assess model performance, accuracy, and robustness. • Detect any drift or degradation in model quality.	• Continuously monitor model behavior in production. • Set up alerts for anomalies, unexpected outputs, or performance deviations. • Address issues promptly to maintain model reliability.	• Evaluate AI models for fairness, bias, and ethical implications. • Mitigate any unintended consequences or discriminatory outcomes.	• Keep track of model lineage, including data sources, preprocessing steps, and training details. • Transparently communicate model decisions and limitations to stakeholders.

ORIGIN

A simple way to start

Data's value is to answer questions to inform decisions and understand outcomes.



What are the hardest questions for you to answer?

What is the mix of "gut-feel" and data in decision-making?

How do you know if you made the right decision or not?

ORIGIN®

Thank you!

Nick White
Head of Decision Science & Experience
nwhite@origindigital.com
origindigital.com