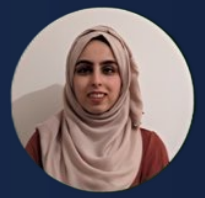
Beyond Hype: The Practical Evolution of Data Science with Generative AI

Data science has always embraced innovation, but the emergence of generative AI presents a truly transformative shift.

Omer Ali Omer
Platform Engineer
@IBM



Narmeen Syeda
Advisory Data Scientist
@IBM

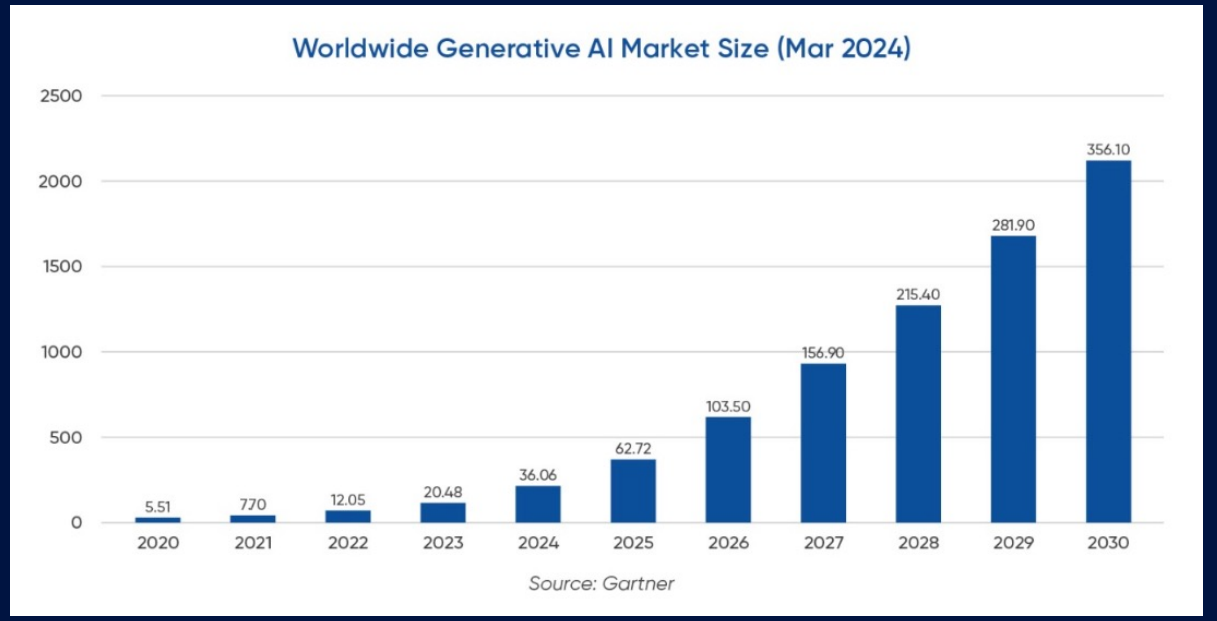
CONF42

Agenda

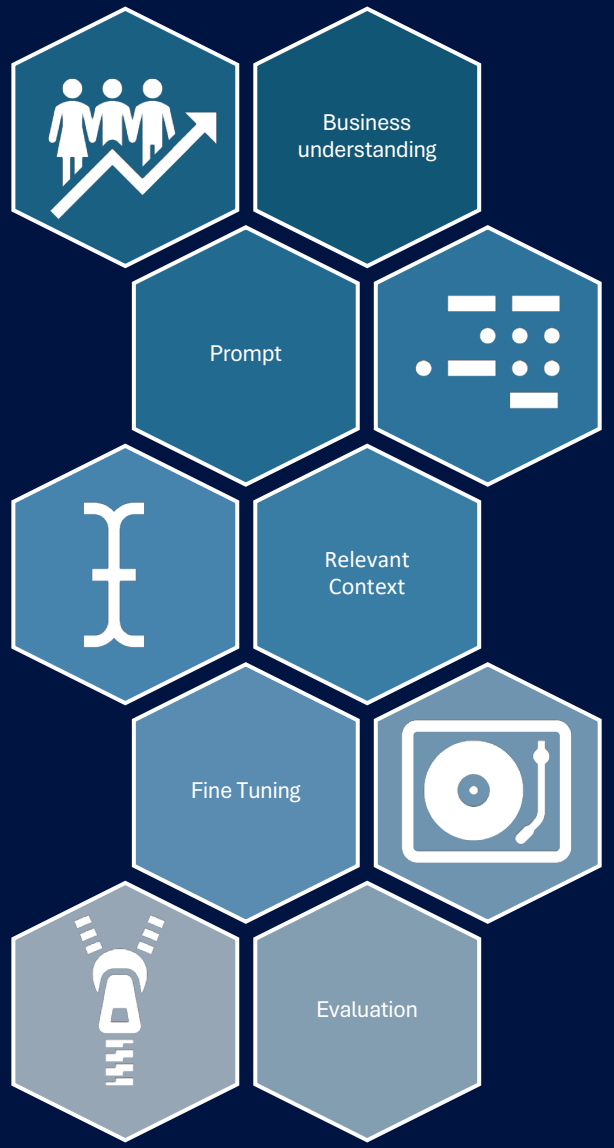
- The Evolving Landscape of Data Science
- Practical Experiences with Generative AI
- The Impact of Generative AI on Data Scientists
- The Future of Data Science with Generative AI
- Next Steps

The Rise of Generative AI

- Technological breakthroughs
- Generative AI creates new content rather than just analysing existing data - it can write text, generate images, compose music, and create code that never existed before.
- Models like GPT-4.5 can generate synthetic datasets that preserve statistical properties while protecting privacy, or create code to implement complex algorithms from natural language descriptions.



From Prediction to Creation: The Paradigm Shift in AI Workflows



Data Scientists are Concerned

- Generative AI rapidly analyses data and produces actionable insights.
- Eliminates the need for building and training machine learning models in some cases.
- Automates lot of data science tasks thus raising concerns.
- The ‘Doom Loop’: AI Will Be Taking Your Jobs In 2024, Leaders Say (Robinson, 2024).

GenAI & Data Scientists: A Dynamic Duo for Innovation

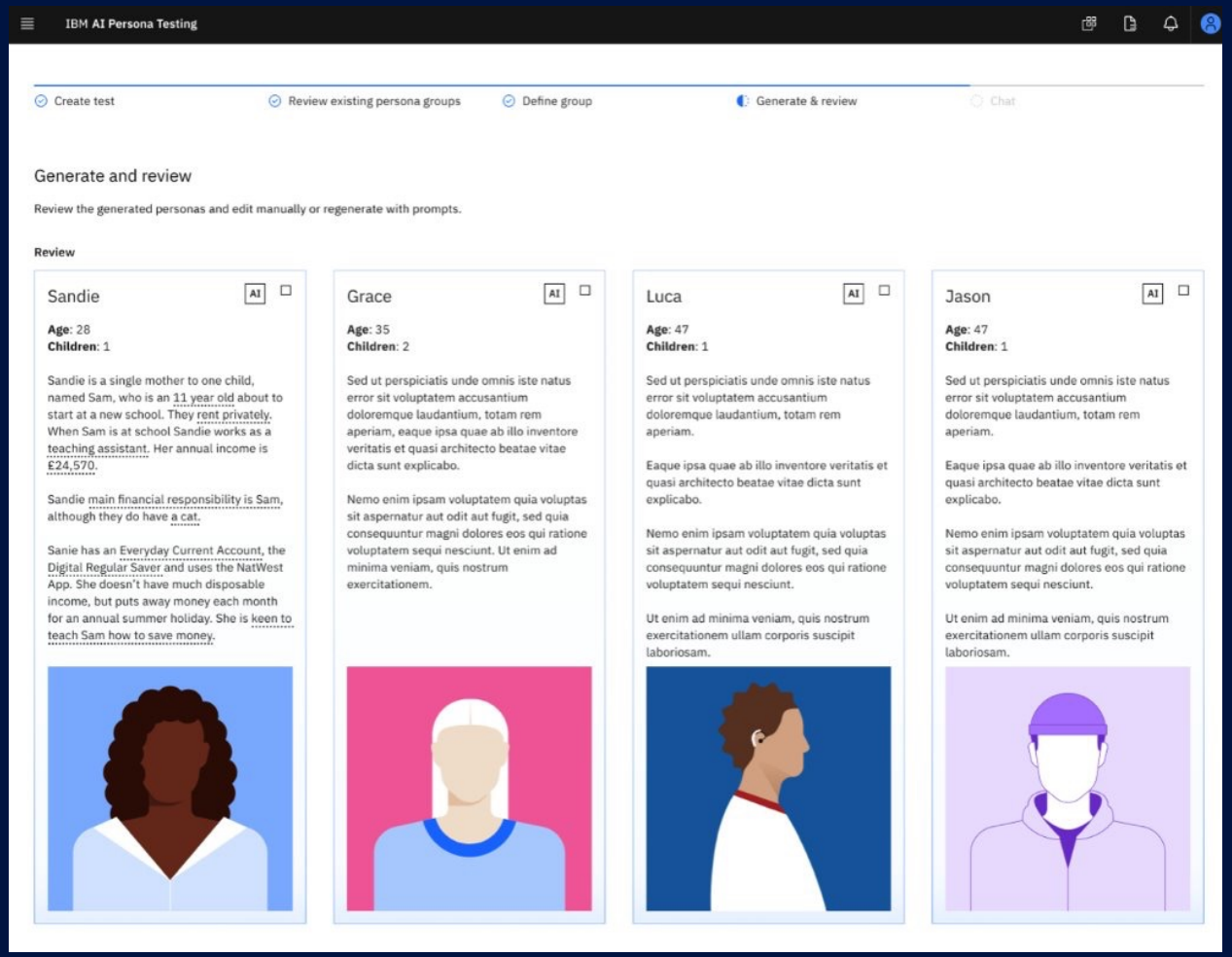
GEN AI CAN INCREASE PRODUCTIVITY BY AUTOMATING MUNDANE TASKS

ACCELERATE TRADITIONAL ML DEVELOPMENT

QUICK PROTOTYPE TO BRING IDEAS TO LIFE SOONER

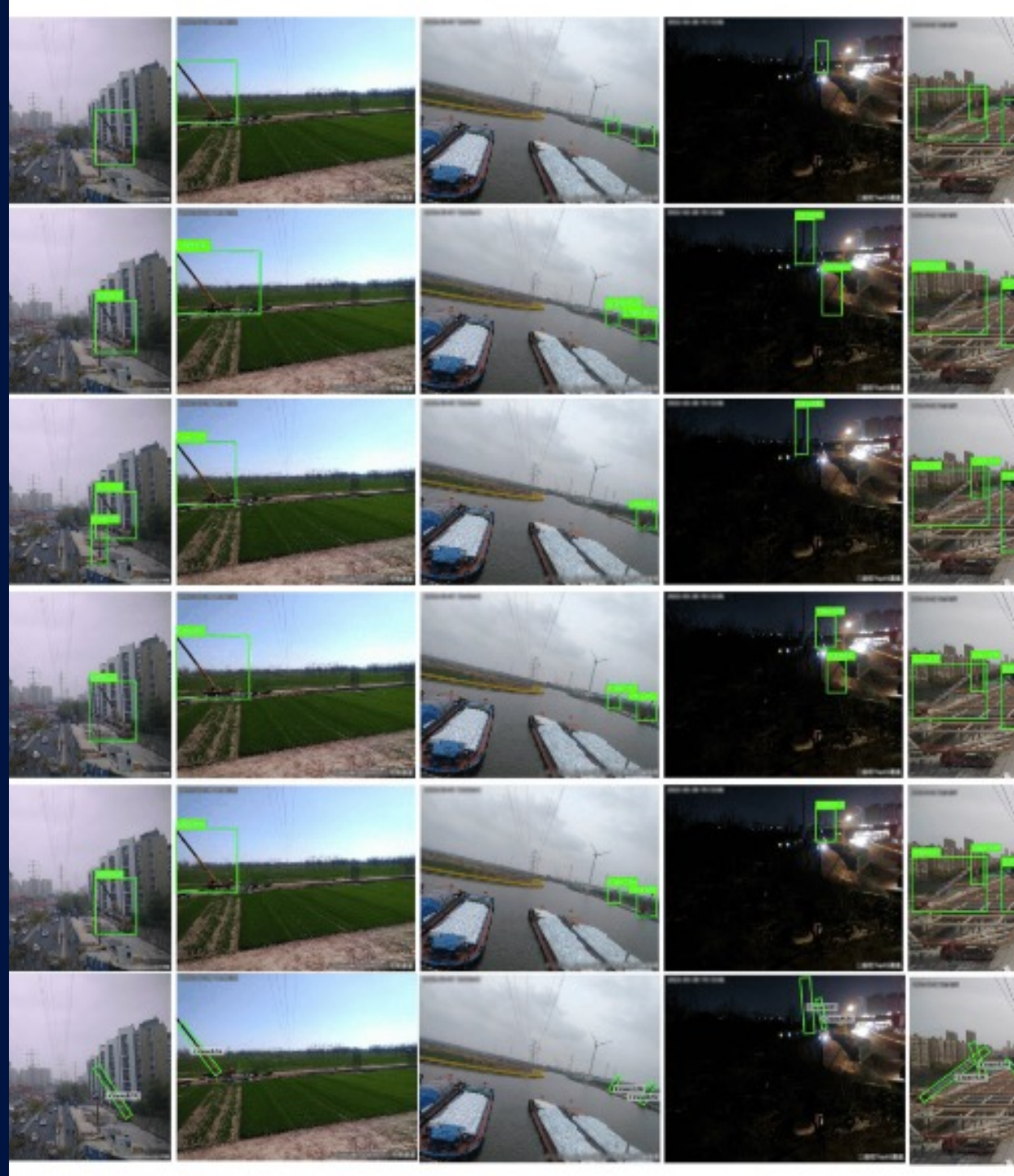
Creating Synthetic Data: Expanding Possibilities

- Finance Services wanted to test marketing strategy with different focus groups.
- Granite model to synthesise customer personas
- Artificially created data that mirrors real-world data.
- Synthesized data to protect *privacy and* overcome *data scarcity*.
- Enables model training when access to diverse data is limited.



Enhancing Model Robustness with Data Augmentation

- Generative AI modifies existing data to increase diversity and improve model robustness.
- Reducing model overfitting
- E.g. Crane Lifting
- Maximo Visual inspection



Leveraging GenAI for Domain Understanding

Foundational understanding of domain

Simplifying Complex Concepts

Analyse data and identify underlying patterns

Summarise large volumes of domain-specific literature

Context- driven guidance

Basic level guidance

Ethics of Generative AI

- Rewrite legal clauses and help in building contracts specific to a client.
- Address potential risks of hallucinations.
- Incorporate human feedback to improve GenAI outputs iteratively

We Still Need Data Scientists!

Gen AI is here to elevate, not replace

Human Oversight and Expertise

Human Creativity and Insight

What’s Next? Emerging Trends in Generative AI

- Multimodal Models
- Reinforcement Learning *in Generative AI*
- Open-Source

Collaborate with AI, Not Against It

Invest in Skills

Embrace Generative AI

THANK YOU