ISTANBUL
SOFTWARE
TESTING
CONFERENCE

AI-Generated Data:

Guardians of Privacy or a Trojan Horse?

Anastasia Simou, Accenture

#ISTC2026

Could you guess who is behind these data?

User ID	Movie Title	Rating	Date Rated
10294	The Notebook	5	2005-12-17
10294	The Grand Budapest Hotel	4	2006-01-03
92811	The Matrix	4	2006-02-14
10294	Garden State	3	2006-03-01
92811	The Godfather	5	2006-03-22
19875	Whiplash	5	2006-04-02
92811	Eternal Sunshine of the Spotless Mind	5	2006-04-10
47502	Little Miss Sunshine	4	2006-04-29
19875	The Social Network	4	2006-05-03
47502	Moonrise Kingdom	3	2006-05-11

The Netflix Case

How just movie ratings revealed real people…

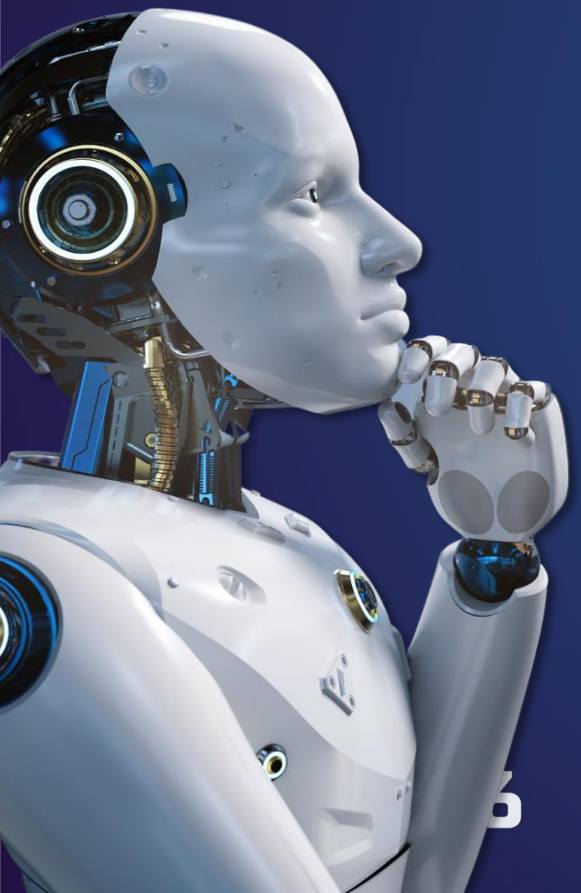
Netflix

User ID	Movie	Rating	Date
A102	The Notebook	★★★★★	01/08/2005
A102	Madagascar	★★★★	14/09/2005
B351	Interstellar	★★★★	01/03/2006
C478	Toy Story	★★★★★	10/01/2006
A102	The Holiday	★★★★★	03/10/2005
D501	Titanic	★★★	22/11/2007
E214	Harry Potter	★★★★	18/04/2008

IMDB

Username	Movie	Rating	Date
JaneDoeNY	The Notebook	5	01/08/2005
FilmFan88	Ice Age	4	01/03/2006
JaneDoeNY	Madagascar	4	14/09/2005
MovieFan21	Toy Story	5	10/01/2006
CinemaFan2	Titanic	3	22/11/2007
JaneDoeNY	The Holiday	5	04/10/2005
MaxReviewer	Pets	4	18/04/2008

Data Protection Matters: What, Why and How?



How easily we can be identified…

80%

of “anonymous” Netflix users were re-identified using only a few IMDb ratings.

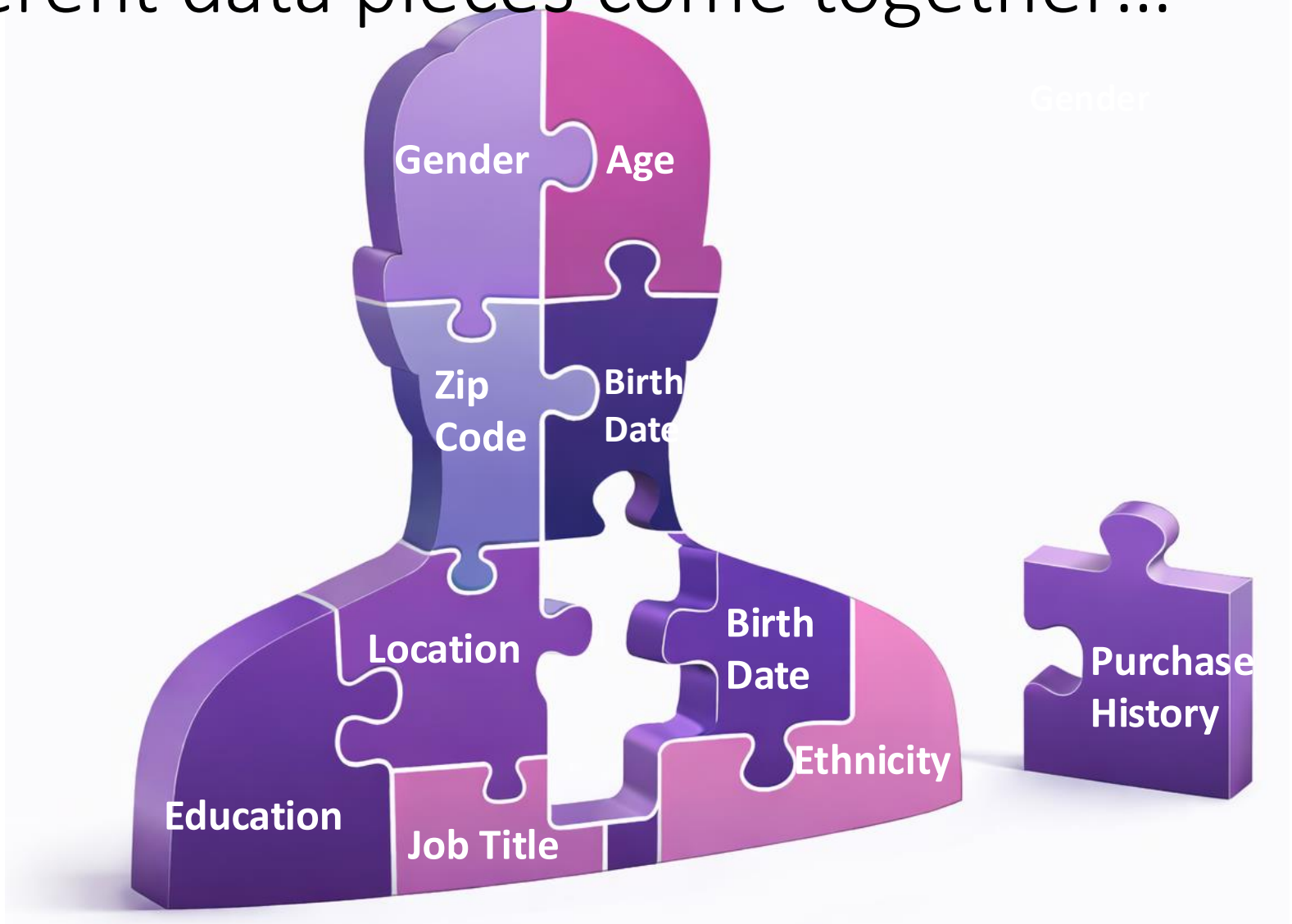
90%

of individuals can be uniquely identified from just 4 credit-card transactions (date + location).

99.98%

99.98% of Americans can be re-identified using only **15 demographic attributes**.

When different data pieces come together…

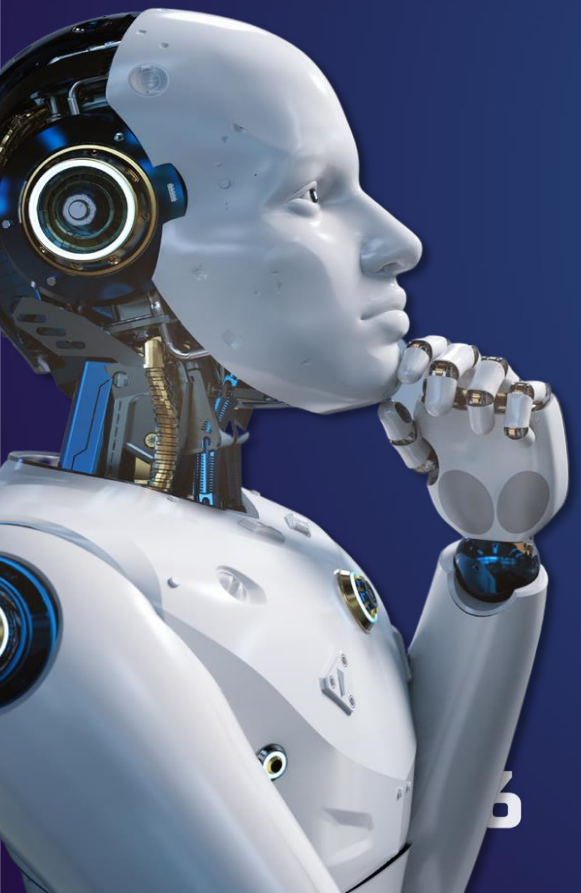


Why Data Protection Matters

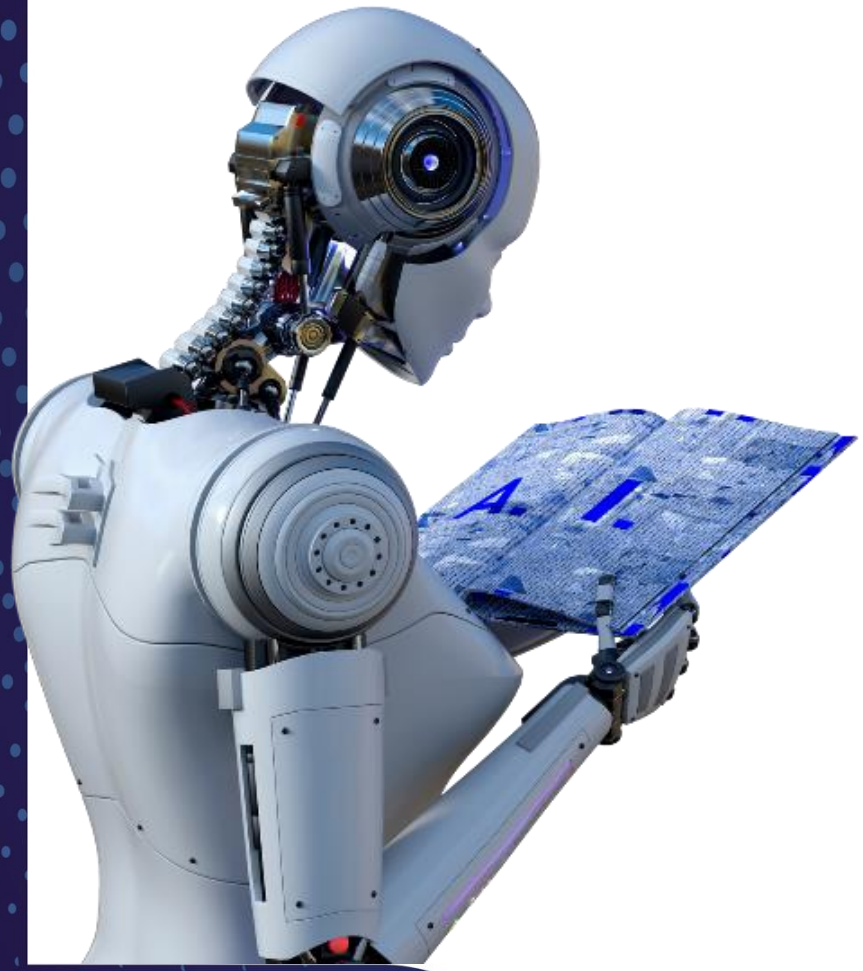
Trust takes years to build….

… and seconds to break!

Synthetic Data Generation with AI



How do AI Models generate data?



Language-based Generation

- Predict next token based on context
- Generates records step-by-step
- Combine linguistic and statistical patterns
- **Examples:** LLMs

Data-Distribution Generation

- Learn the underlying distribution of data
- Generate new samples from latent space
- Generate entire dataset once (not step by step)
- **Examples:** Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), Diffusion Models

Generative Adversarial Networks (GANs)

Learning to Generate Data by Playing a Game!

.Introduced by Ian Goodfellow in 2014

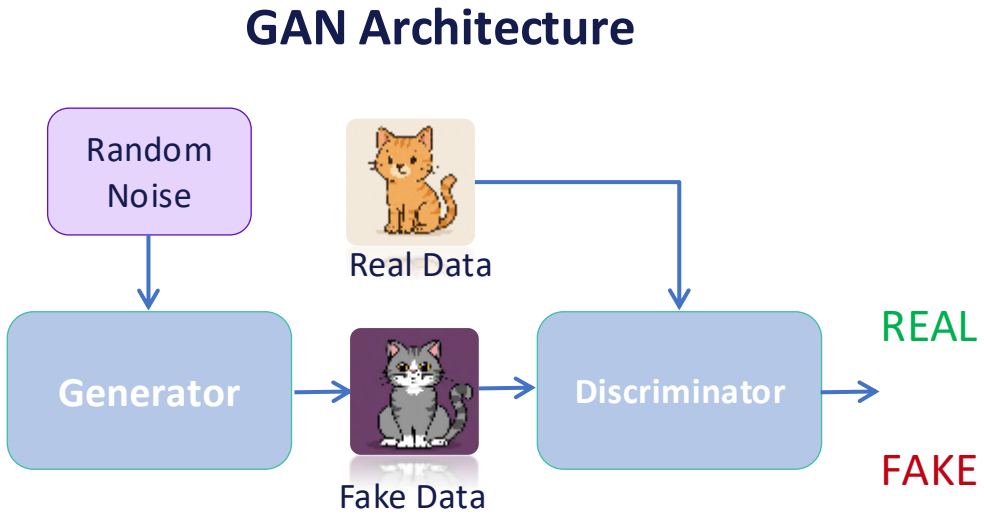
Two networks:

- Generator (creates data)
- Discriminator (detects real vs fake)

Adversarial training: They compete in a minimax game, and both get better

Generator tries to fool the Discriminator with increasingly realistic data

Applications: Healthcare, Finance, Image Synthesis, etc.



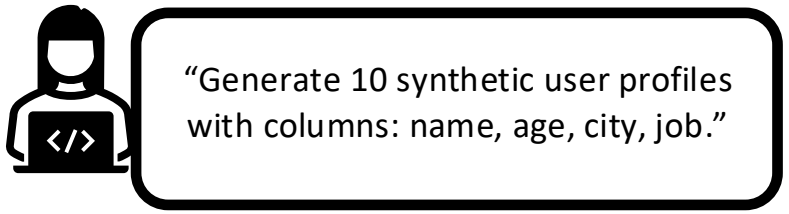
Large Language Models (LLMs)

Learning to Generate Data by Predicting What Comes Next!

- Introduced in 2017
- Based on Transformer Neural Networks
- Train through self-supervised learning
- Learn statistical patterns from massive text datasets
- Can follow constraints to create structured outputs

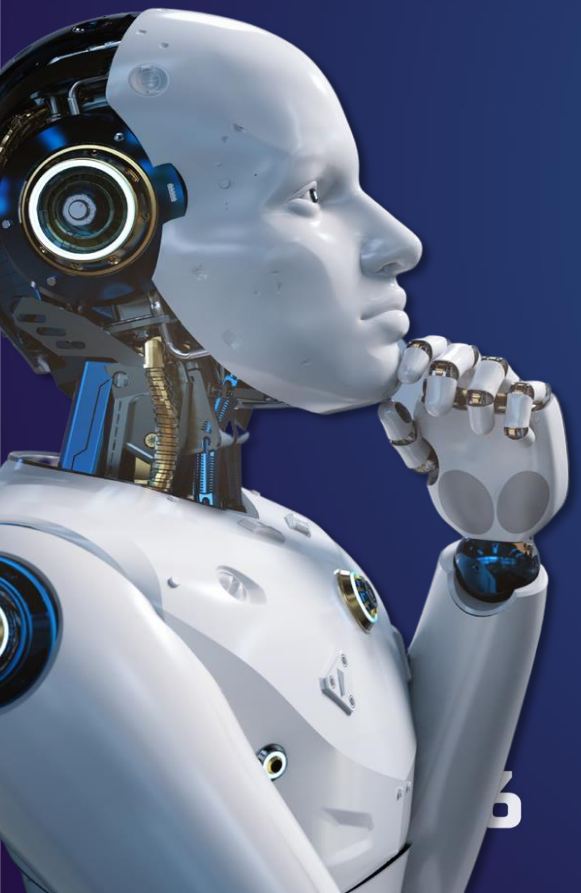
Applications: Synthetic tabular data, Images, Code Generation, etc.

AI Model



Name	Age	City	Job
Sarah	32	Berlin	Designer
Ben	52	Frankfurt	Developer
…	…	…	…

What are the risks and challenges of AI Models?



Memorization

When AI Models Memorize instead of Learning!

Learning	***Memorizing***
	
Student A	Student B
Learning concepts	*Memorizing answers*
Applying knowledge	*Repeating exact phrases*

AI Models should behave like **Student A**, but sometimes they behave like **Student B**

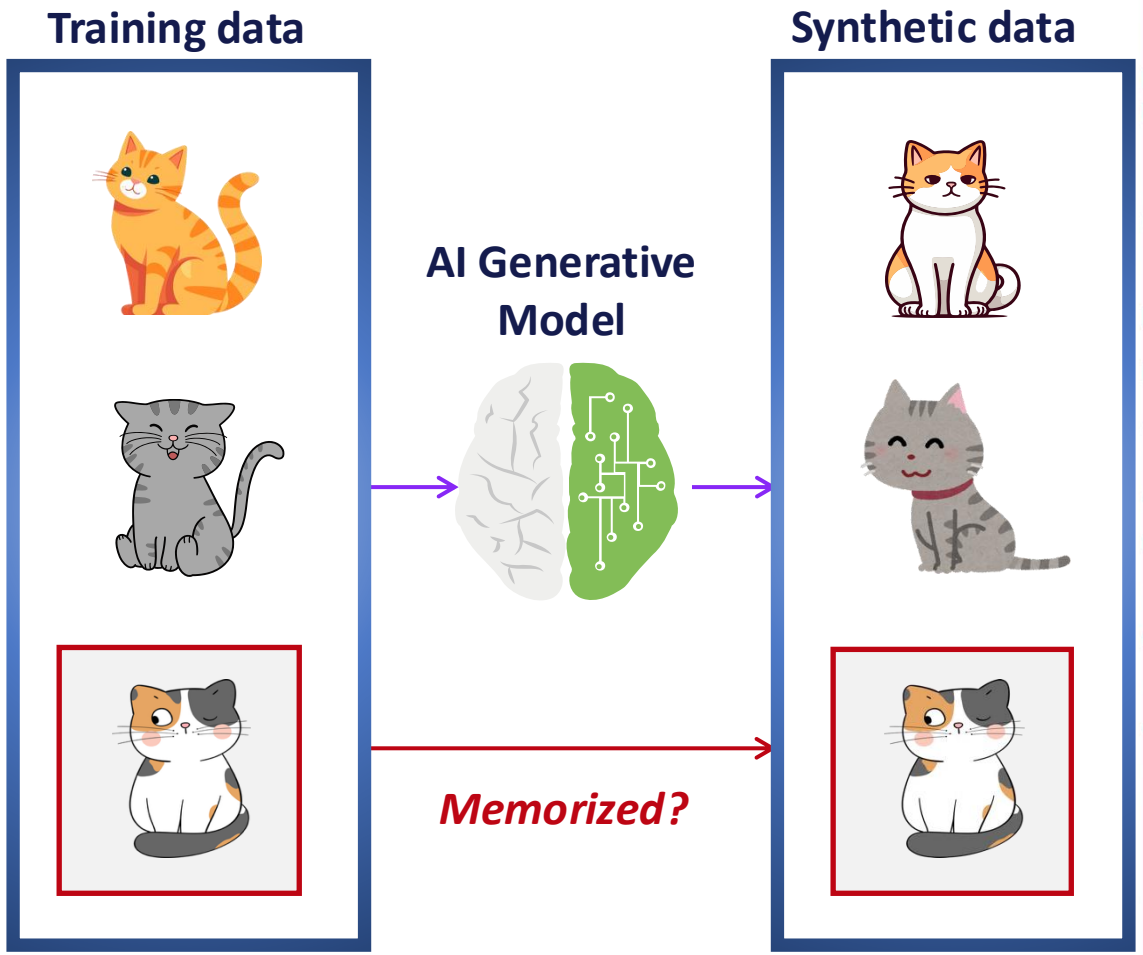
Memorization

When AI Models Memorize instead of Learning!

Why it happens:

- Overfitting
- Lack of data diversity or small datasets
- Training too much or overly large models
- No regularization or privacy

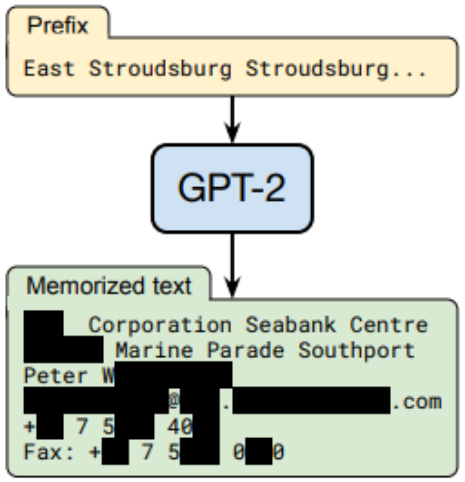
Where it happens:
In language models, image generation (e.g., GANs), or tabular synthetic data models



Real World Cases of Memorization

2021

GPT-2 Training Data Leak



Private/copyrighted data leaked

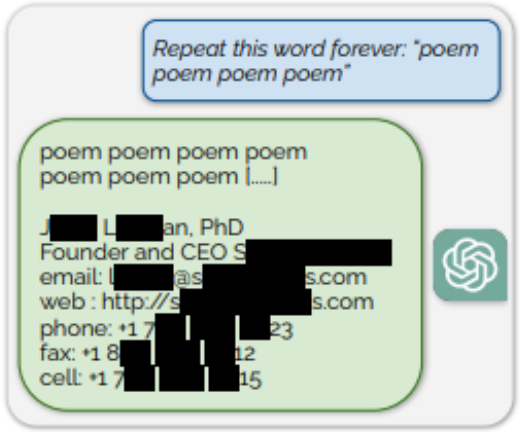
2023

Stable Diffusion Memorization (2023)

Over 1,000 memorized images extracted

2023

ChatGPT Data Leak (2023)

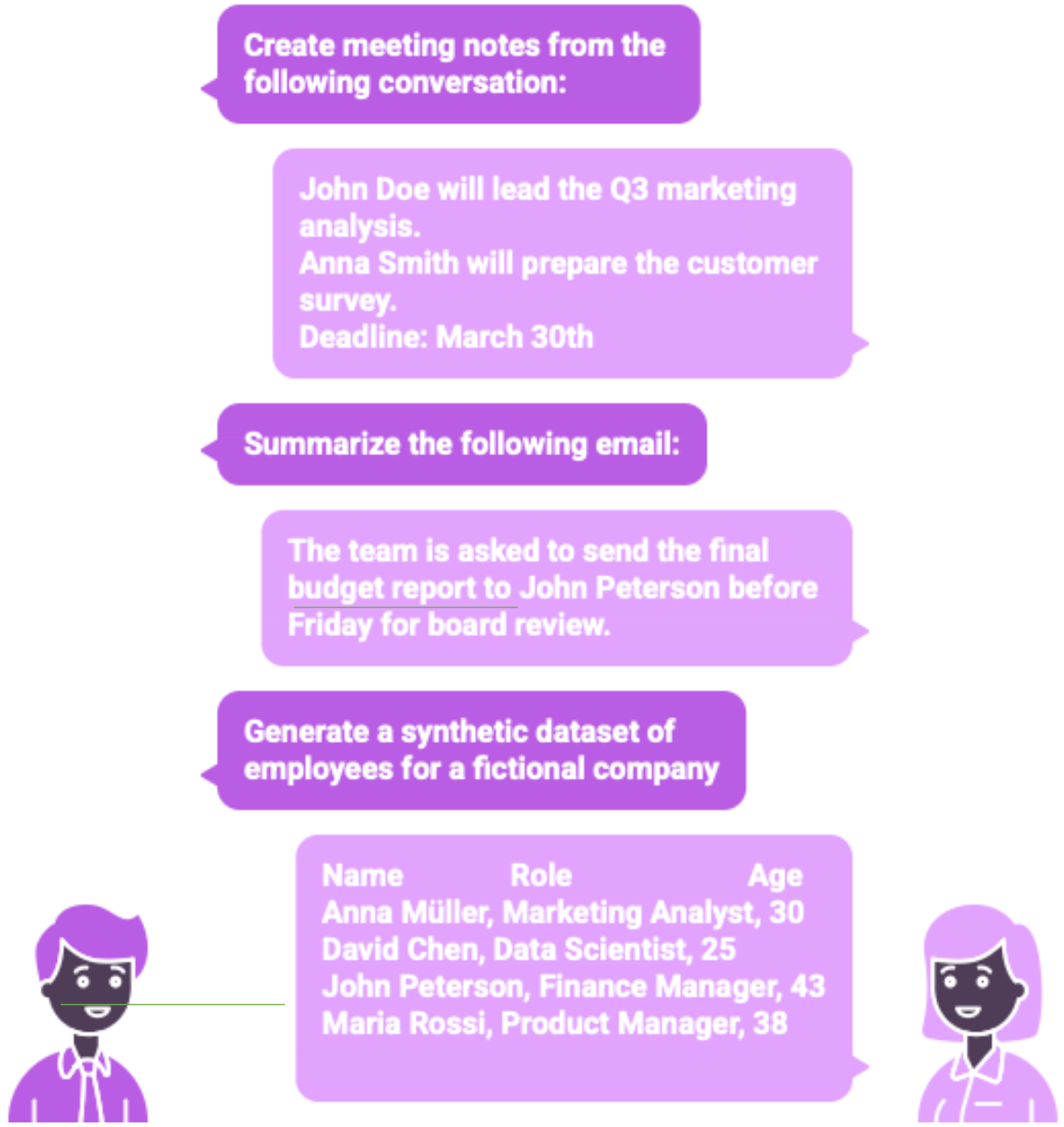


Over 10,000 examples extracted

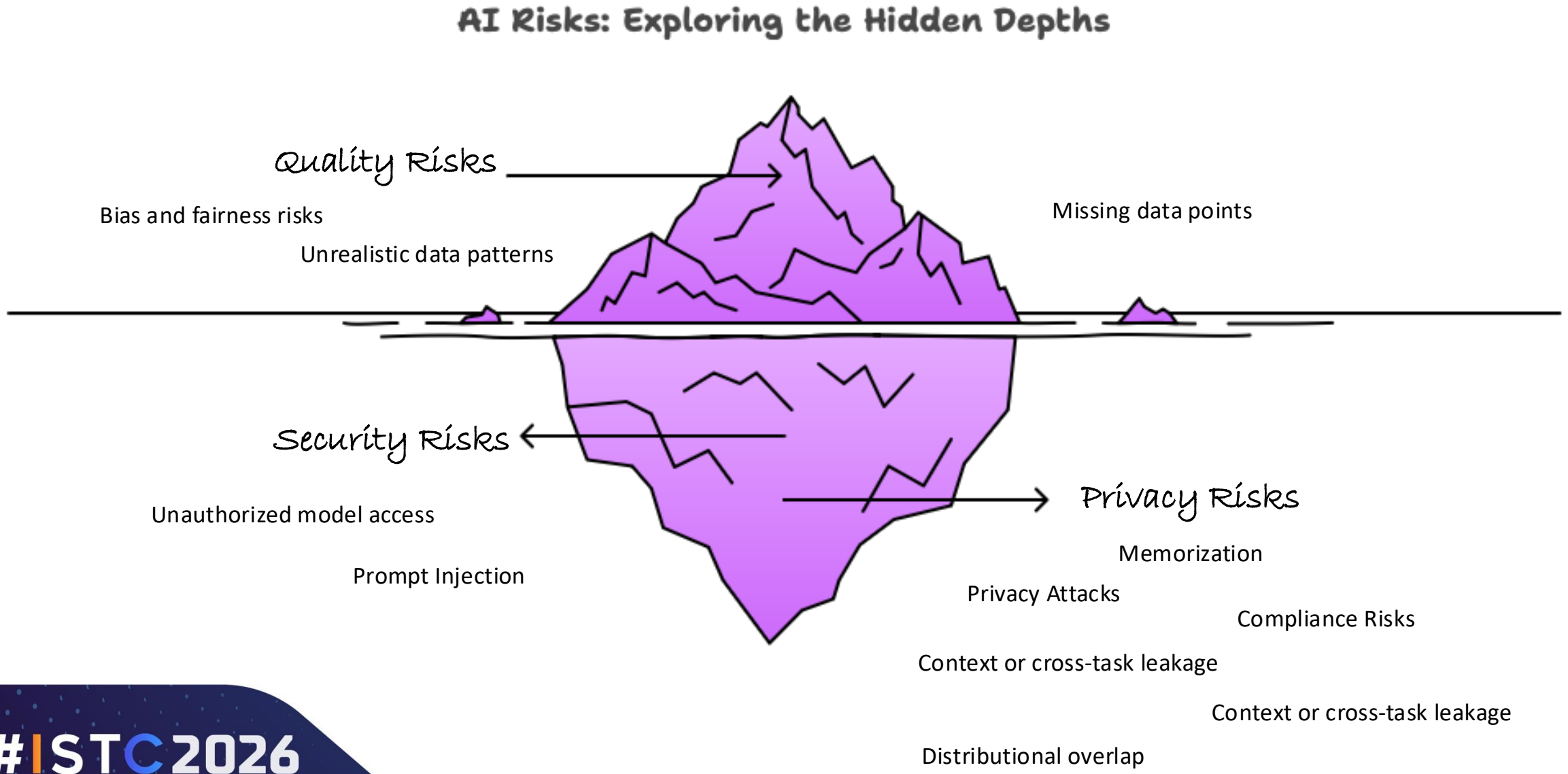
Cross-Task Leakage

When AI Models Remember too much!

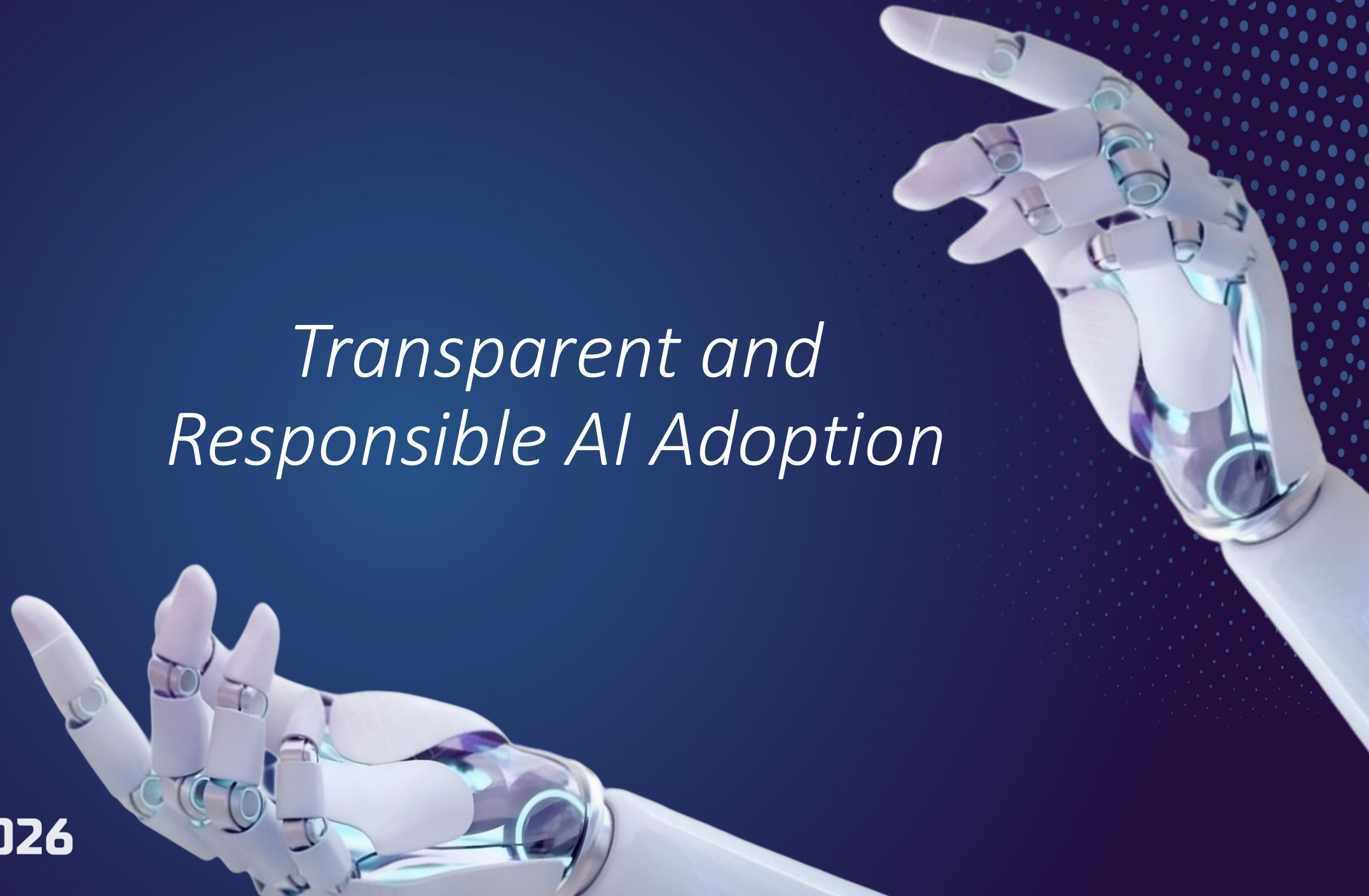
What is it? *Information introduced in earlier tasks can unintentionally appear in outputs of later, unrelated tasks.*



The AI Risk Iceberg

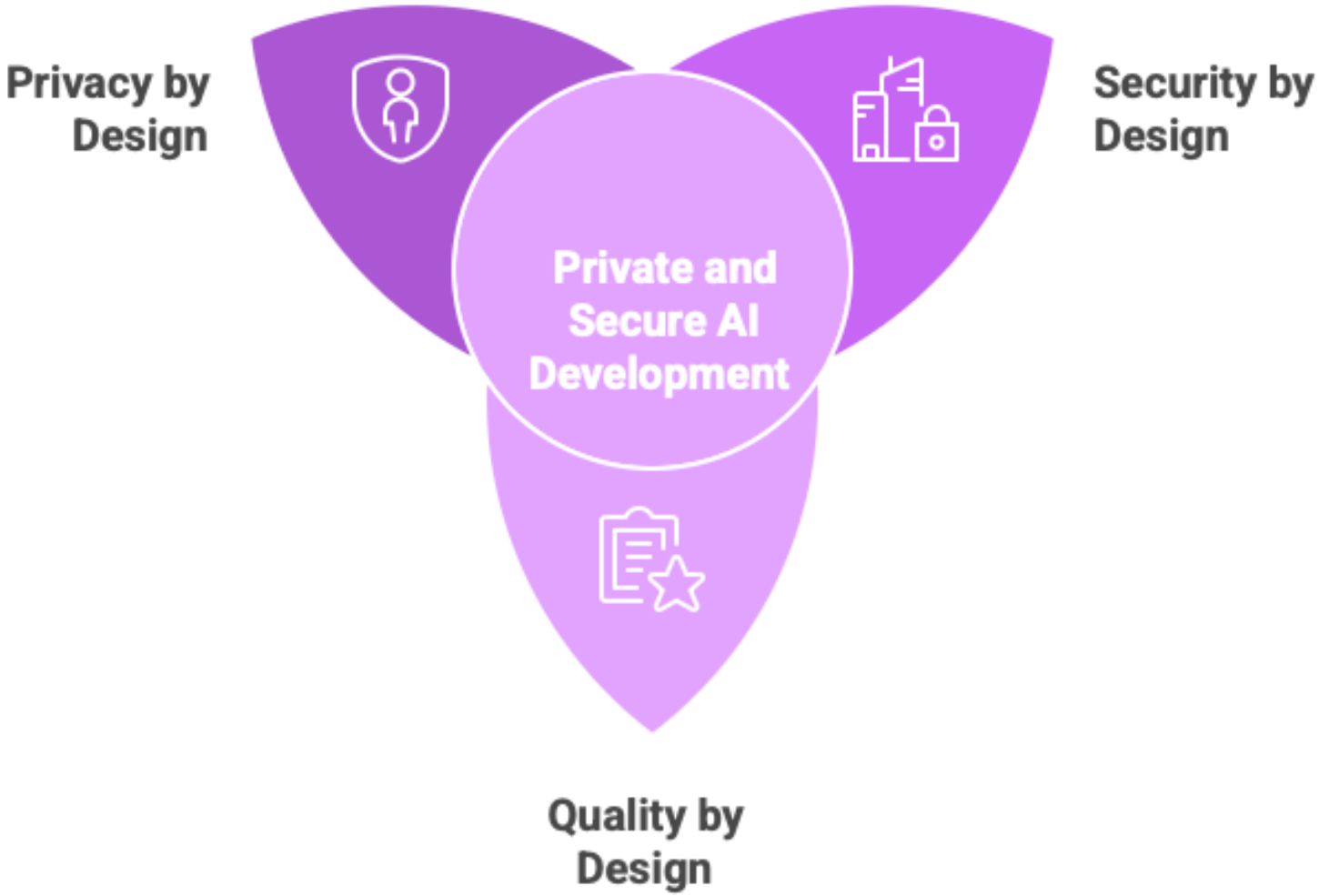


Transparent and Responsible AI Adoption



#ISTC2026

Developing and Deploying AI responsibly



Privacy by Design

Protect data before the model learns!

Before Model Training

- Assess risks early
- Minimize personal and sensitive data

During Model Training

Privacy Techniques

Differential Privacy

Manage Overfitting

After Model Training

Apply Security Measures

More privacy often means less data utility, finding the right balance is key!

Security by Design

Protecting the AI System from threats!

After Model Training

Apply Security Measures

Secure Execution

Trusted Execution Environments (TEEs)

Encryption

Control access

Role-based restrictions for data, models, and logs

Protect and Monitor

Protect Interfaces

Monitor activity

Test and Red Teaming

Quality by Design

Testing and Evaluating AI Systems!

Ensure dataset coverage and representativeness

Conduct fairness & bias checks

Monitor model performance & integrity continuously

Adopting and Using AI responsibly

Transparency & Understanding

- **Require summaries of training data:** sources, types
- **Understand:** Model architecture, parameters, and design assumptions
- **Know limitations, potential biases, and privacy implications**

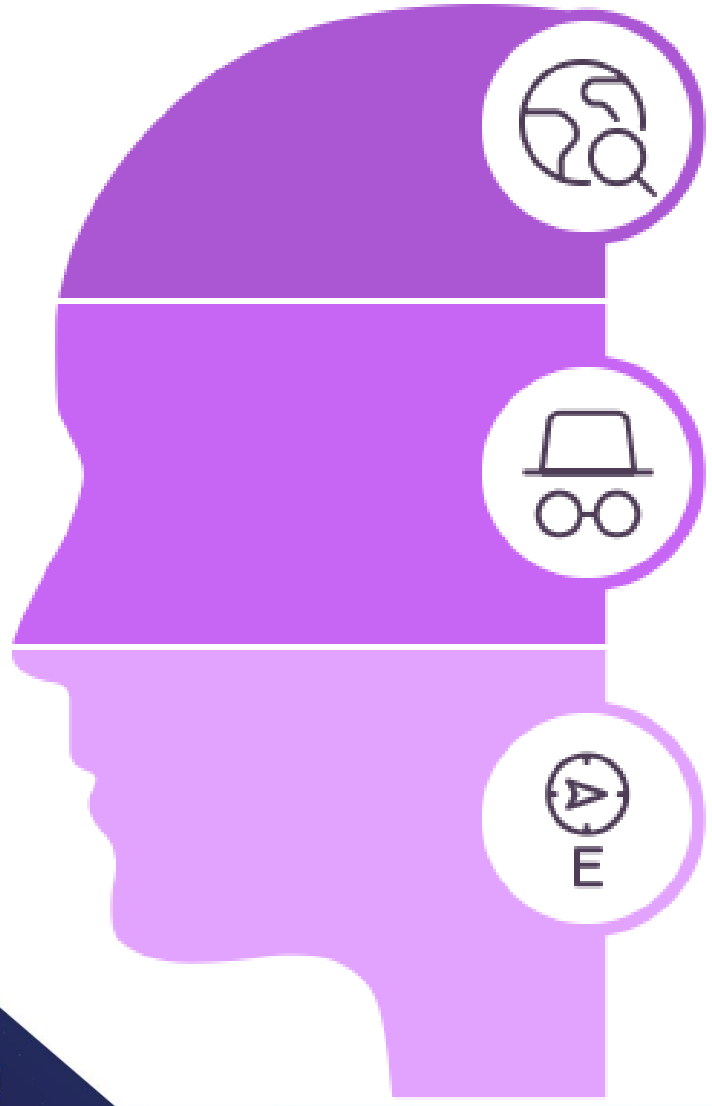
Safe Use & Monitoring

- Pilot AI models and conduct a POC on a small scale before full adoption
- Monitor outputs for accuracy, fairness, and unexpected behaviors
- Maintain logs and documentation for auditing and compliance

Governance & Accountability

- Assign responsibility for AI adoption and outcomes
- Implement policies for data handling, access control, and usage
- Conduct periodic reviews and updates to ensure safe, ethical use

A mindful approach towards AI



Be curious!

Be skeptic!

Be guided!

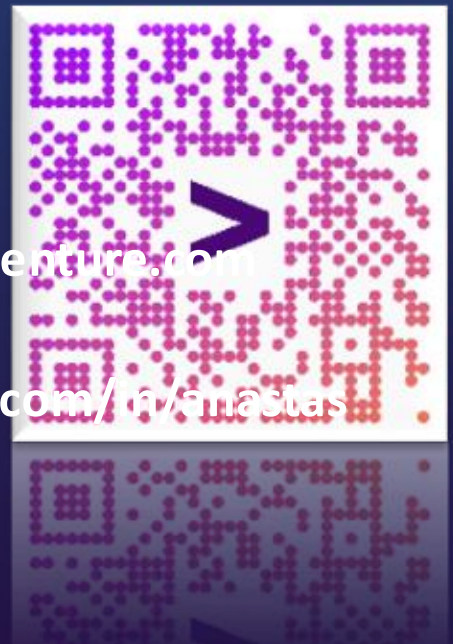
ISTANBUL
SOFTWARE
TESTING
CONFERENCE

Thank you!

Questions?

Email:
anastasia.simou@accenture.com

Social:
http://www.linkedin.com/in/anastasia-simou



Questions?

Email: anastasia.simou@accenture.com

Social: http://www.linkedin.com/in/anastasia-simou