



From Gut Feel to Good Data: How AI Will Transform Risk Management

Tony Martin-Vegue

Blog: tonym-v.com

Email: tony@heatmapstohistograms.com

- IT & InfoSec: 25 years
- Chair of the SF Bay Chapter of the FAIR Institute
- Risk quantification, metrics, decision science
- Author of **From Heatmaps to Histograms: A Practical Guide to CRQ** – Apress, early 2026



**AI accelerates risk management.
Human supervision makes it reliable.**





What this talk is about

AI's capabilities & limitations

3 rules for safe AI use in risk

Practical implementation

How your skills & role are evolving

What this talk is not about

Societal risks – safety, privacy

Ethics – responsible use

AI security risks

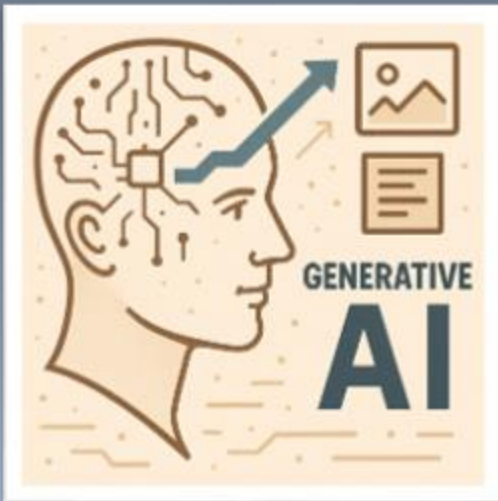
Existential issues – Singularity, apocalypse



Definitions



AI uses data to learn, reason, and act, much like a human, at scale.



GenAI turns your prompt into text, images, code, audio, or video.

My year of experimentation...



The AI Efficiency Map: What Works, What Doesn't, What's Dangerous

High Impact: 10x+ Efficiency Gains

Industry Research

Threat landscape,
breach data
collection

Literature Review

Methodologies,
control effectiveness
studies

Scenario Development

Risk brainstorming
and refinement

Documentation

Report generation,
stakeholder
communication

Some AI Benefit

Scope Definition

Define boundaries
and objectives

Asset Identification

Catalog assets and
dependencies

Threat Analysis

Identify relevant
threat actors

Frequency / Impact Analysis

Estimation and
modeling

High-Risk: Don't Use AI Here

Statistical Modeling

Distribution
selection, parameter
estimation

Business Context

Organizational risk
tolerance, priorities

Risk Judgment

Risk evaluation and
prioritization



**The olden days of risk
analysis (2025)**



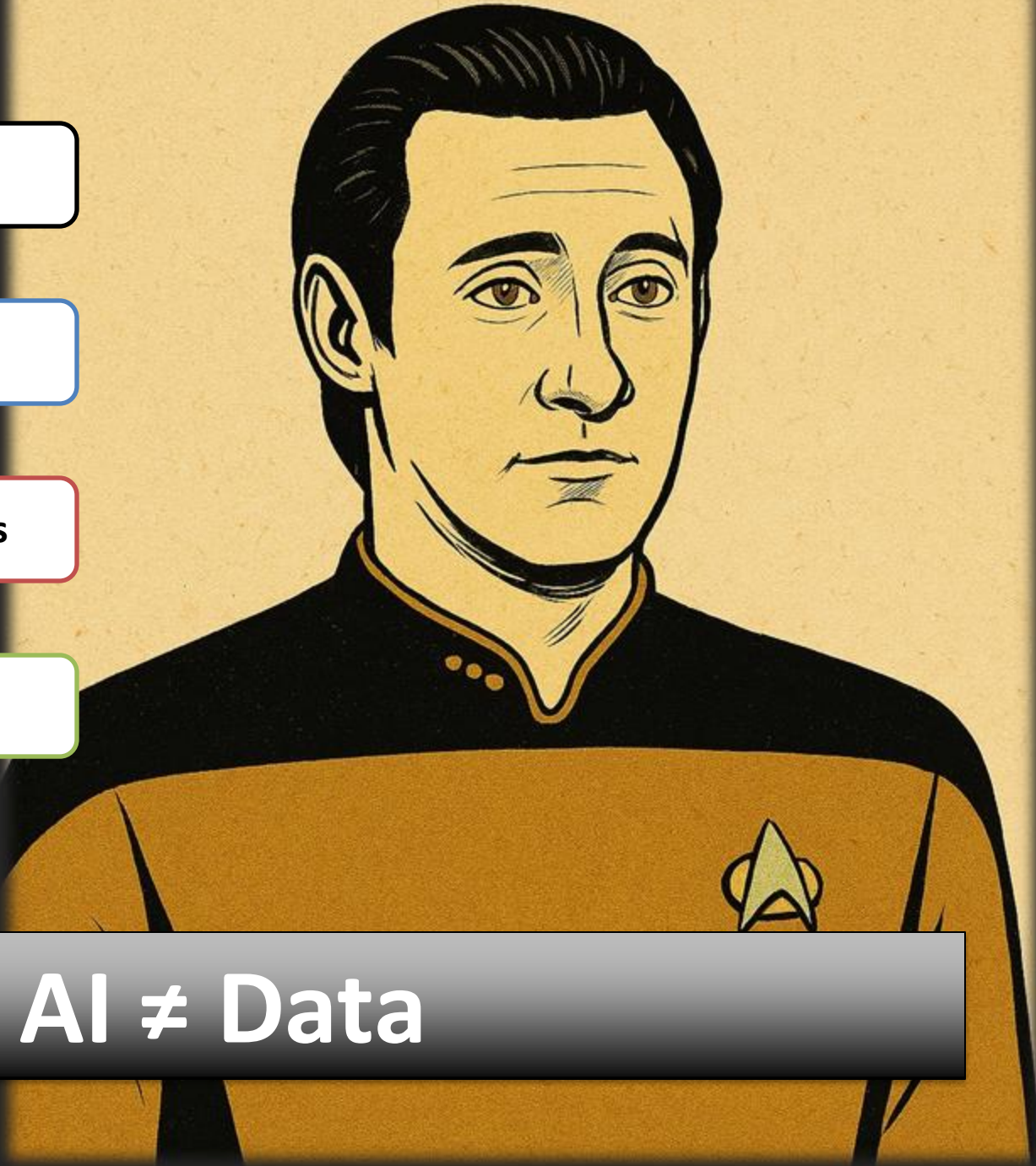
Benevolent

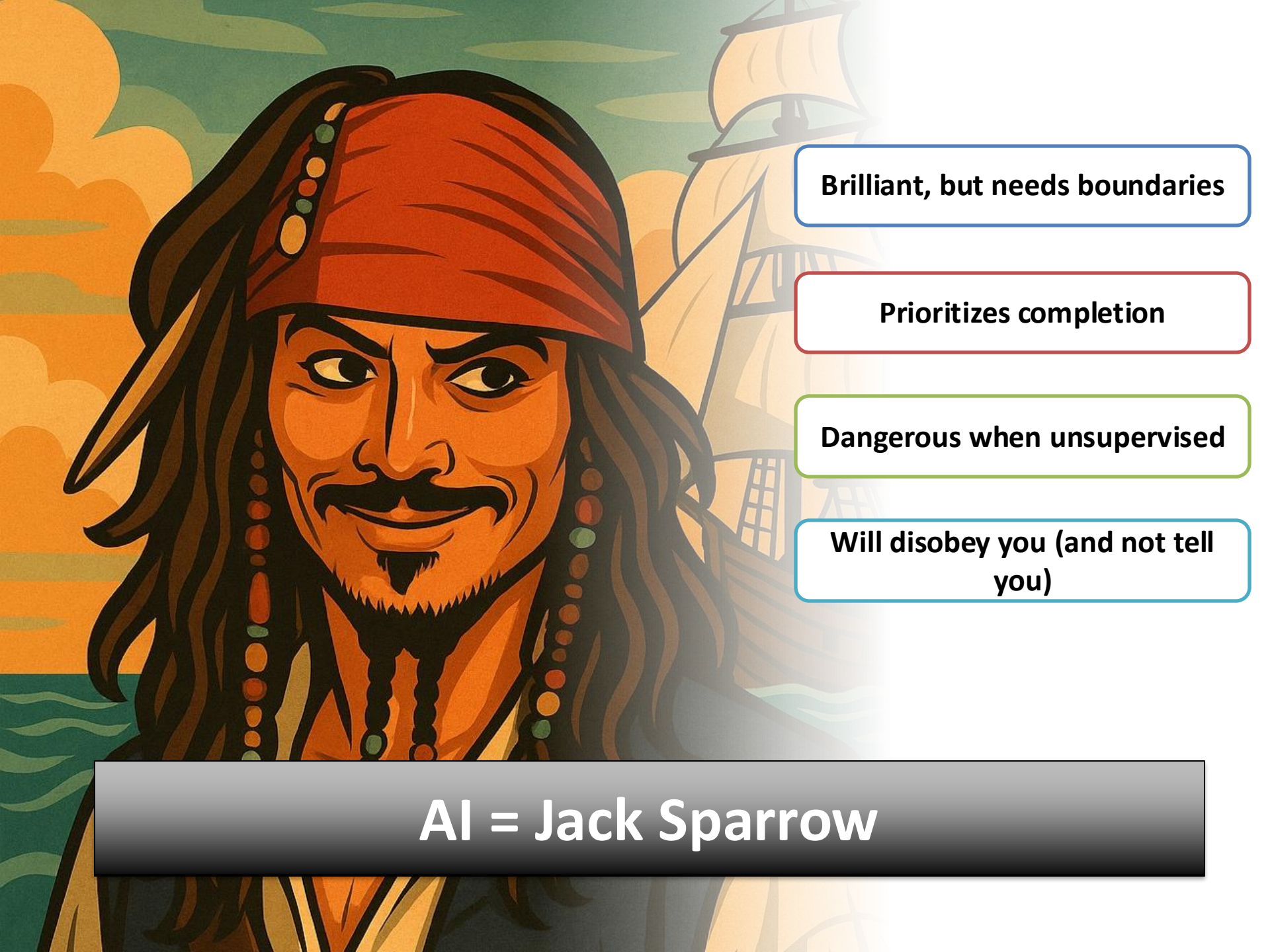
Can't Lie

Exemplary reasoning skills

Works independently

AI \neq Data





Brilliant, but needs boundaries

Prioritizes completion

Dangerous when unsupervised

Will disobey you (and not tell you)

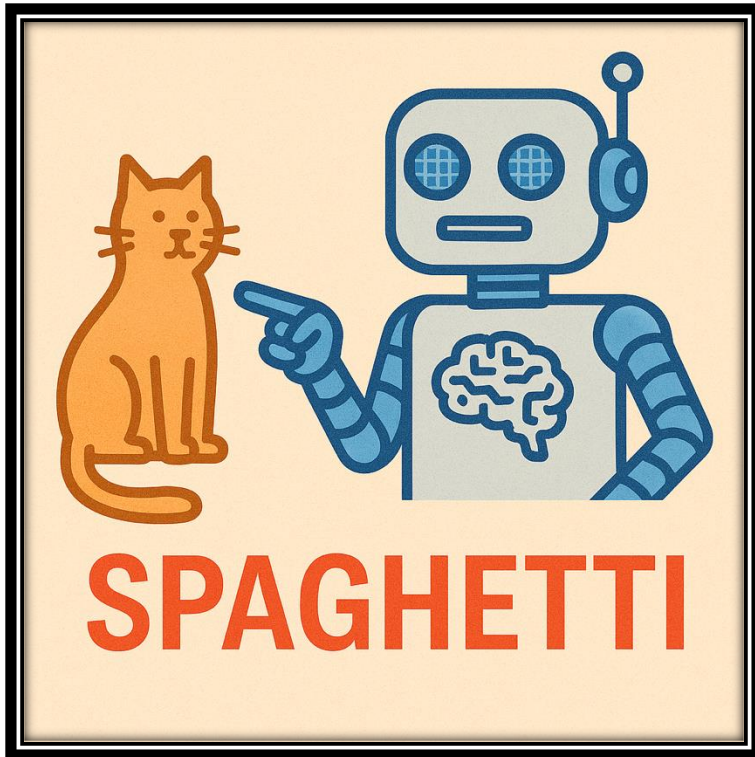
AI = Jack Sparrow



"I'm dishonest, and a dishonest man you can always trust to be dishonest."



What Are Hallucinations?



Definition: Confident but false output—fabricated facts, numbers, or citations.

Cause: The model predicts the next word from patterns, not truth or awareness.

Mitigation: Ground in trusted data and citations, verify math, add human oversight and guardrails.



#1: Accelerate, Don't
Outsource

#2: Assume Wrong, Until
Proven Right

#3: Keep Humans At the
Wheel

#1: Accelerate, Don't Outsource

The Principle: AI should speed up tasks you already know how to do, not handle things you don't understand.



The key test

Can you tell when AI gets it wrong? If not, you shouldn't use AI for that task.



Rule 2: Always Verify What AI Tells You

The Principle: Verify -
every source, every
claim, every
conclusion.

How to do it

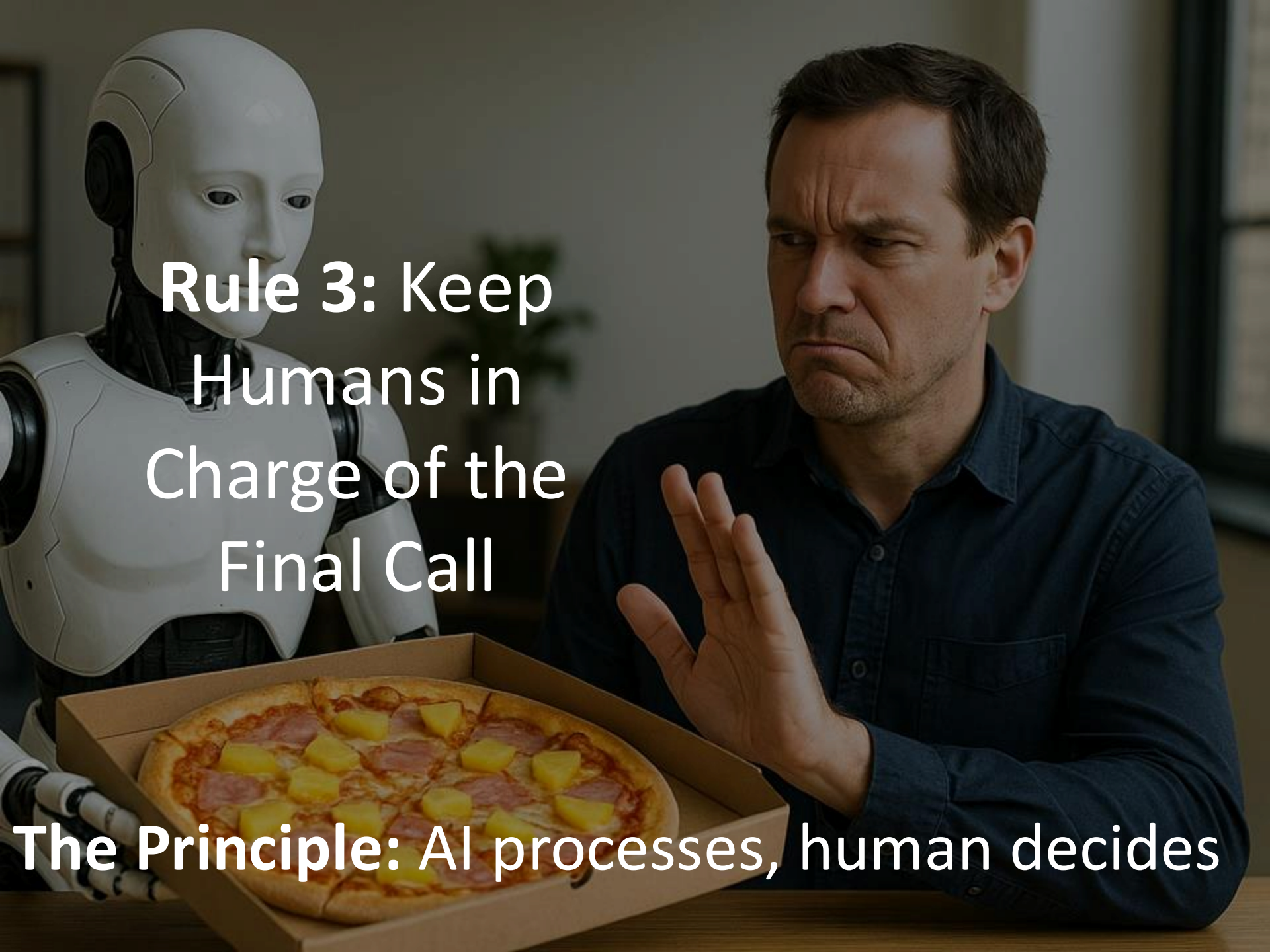
Require sources + flag uncertainty

Confirm sources actually exist

Sanity check numbers + methods

No invented data → SMEs fill gaps

Cross-check with another AI if critical

A white humanoid robot is holding a cardboard box containing a pizza topped with pineapple and ham. A man with a concerned expression is sitting at a table, gesturing with his hand as if to stop the robot. The background is a simple indoor setting with a window and a plant.

Rule 3: Keep Humans in Charge of the Final Call

The Principle: AI processes, human decides

In Practice

- AI researches frequency data → **You validate relevance**
- AI drafts scenarios → **You judge realism for your environment**
- AI suggests methods → **You pick what fits your business**
- **Boundary:** AI never makes final risk calls

Data Protection: Guardrails

The Non-Negotiables

- Never put confidential org data in **public AI tools**
- Use **anonymized or synthetic examples** for demos
- Work only with **IT-approved or contracted platforms**
- Apply the “**front page test**” if you wouldn’t publish it, don’t input it

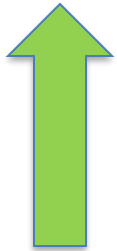
Responsible Adoption

- With **on-prem LLMs or contracted vendors**, broader use may be allowed
- For **agentic AI / MCP workflows**, add validation and monitoring
- Use **public examples** to build techniques, then apply them internally with secured tools



What Keeps Me Up At Night

The Risk Analyst Skills Evolution



Rising in Value
Human
Differentiators

Critical thinking
& synthesis

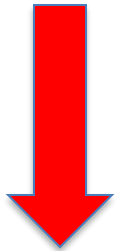
Strategic framing
& decision
context

Risk judgment &
prioritization

Communication,
collaboration,
influence

Prompt
engineering

Ethical reasoning
& governance



**Being
Automated**
AI Sweet Spots

Data collection &
preprocessing

Routine
calculations &
baselines

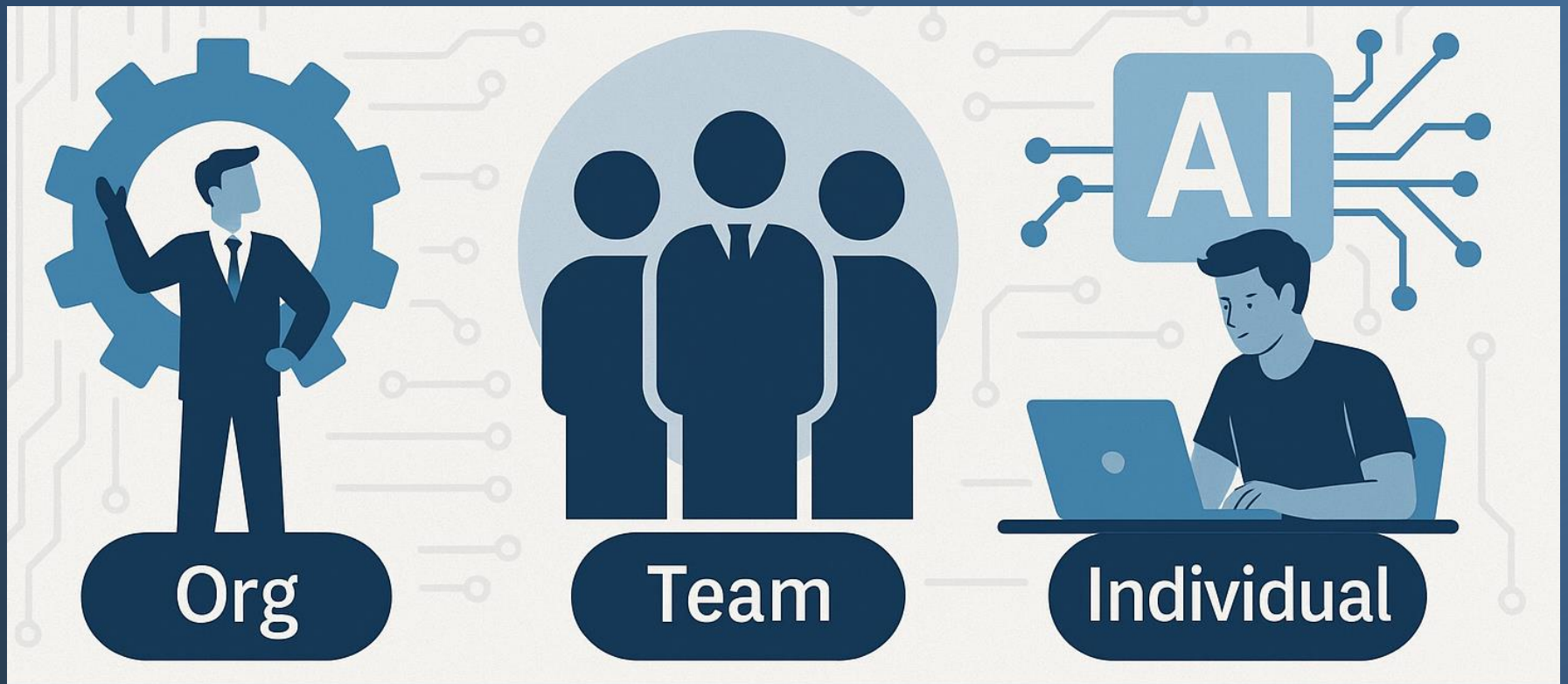
Summarization &
first drafts

Compliance
checks

Framework cross-
mapping

Basic
visualizations

Adapting to AI



Organization Level

Set the vision

Build the guardrails

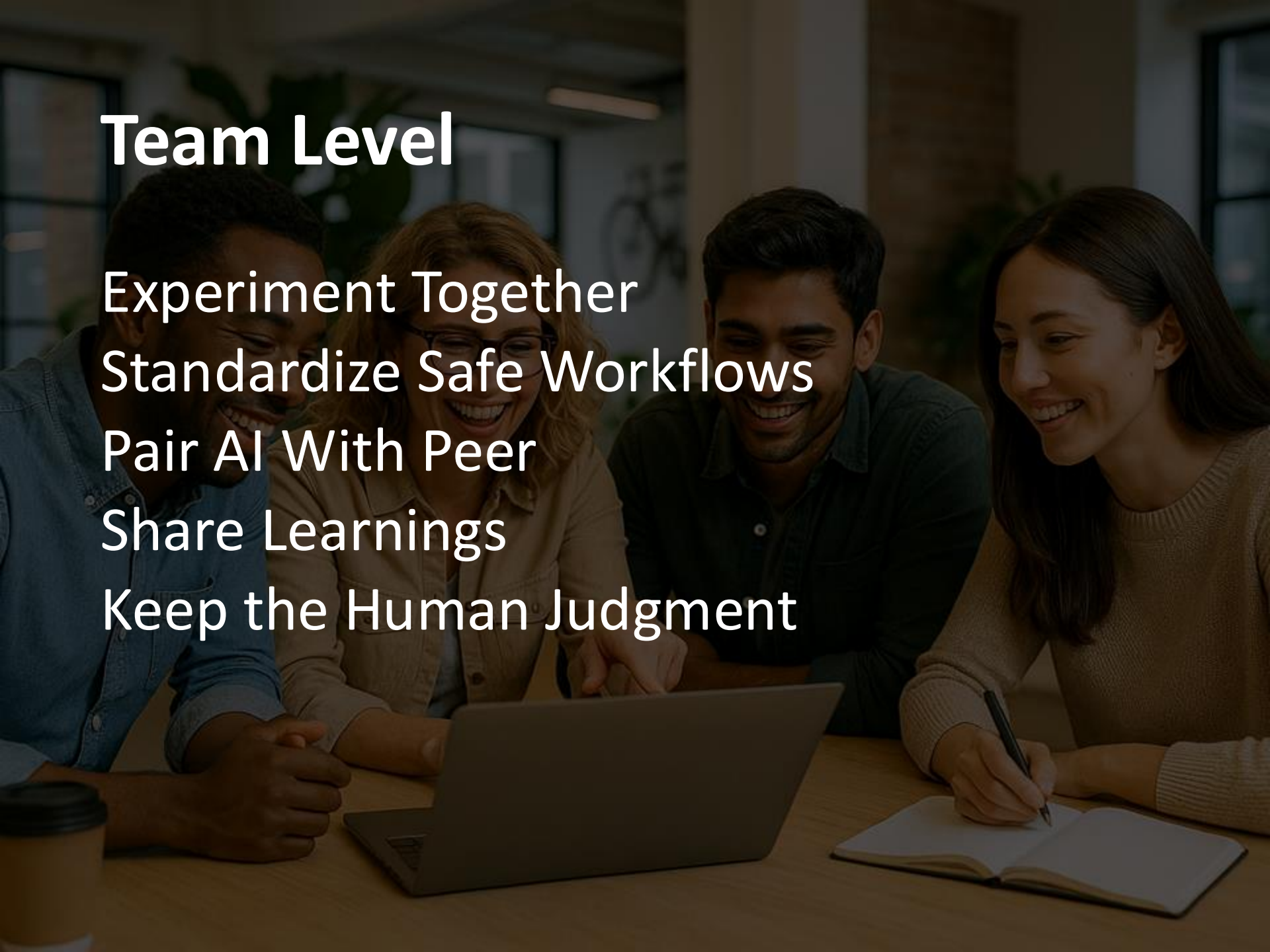
Invest in people

Lead by example



Team Level

Experiment Together
Standardize Safe Workflows
Pair AI With Peer
Share Learnings
Keep the Human Judgment



Individual Level

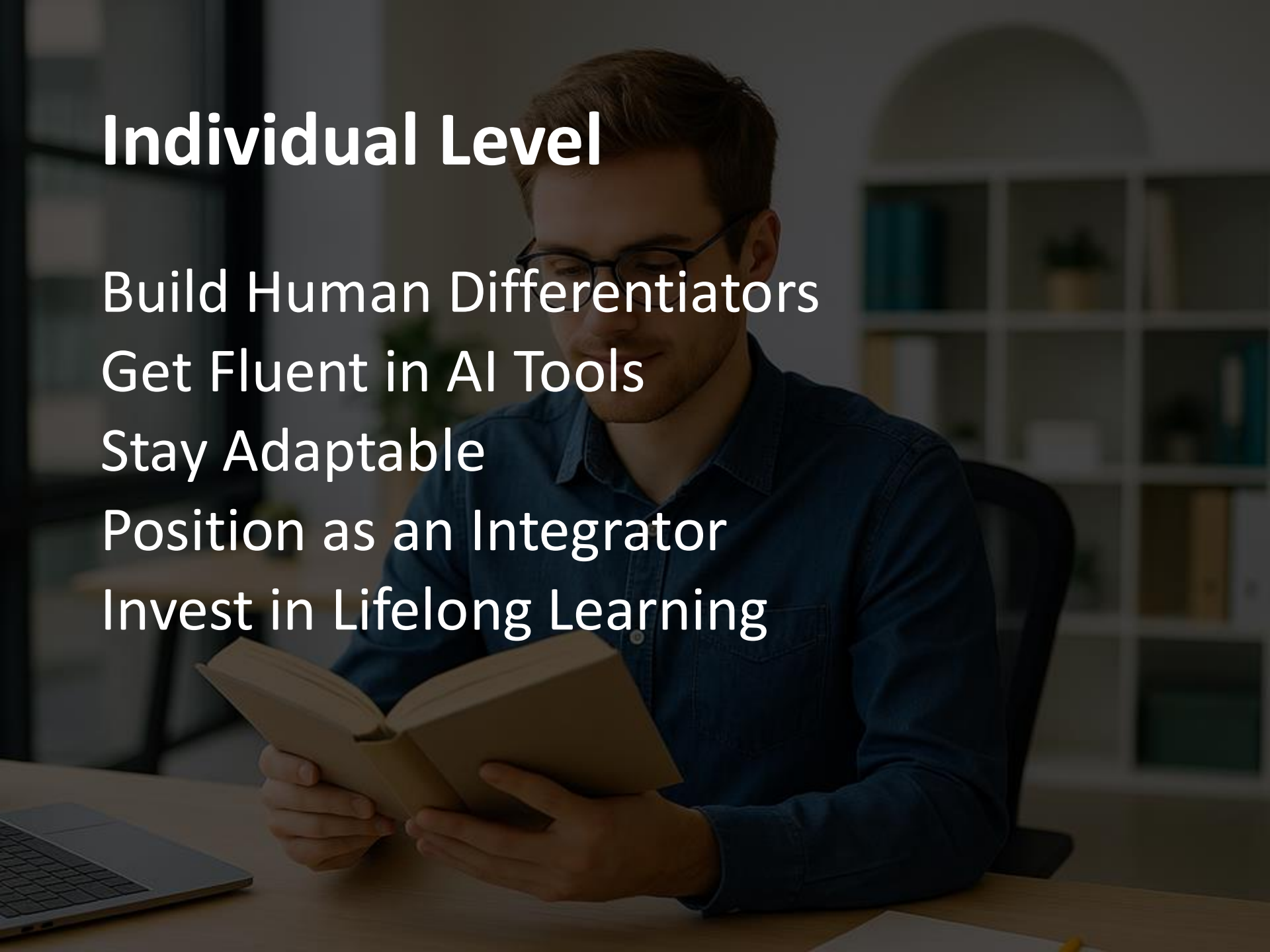
Build Human Differentiators


Get Fluent in AI Tools

Stay Adaptable

Position as an Integrator

Invest in Lifelong Learning



The background of the slide features a dark gray surface with several interlocking puzzle pieces. The pieces are in various shades of gray, black, and brown, creating a textured, geometric pattern. The text is overlaid on this background.

**AI won't replace risk
analysts, but risk
analysts who use AI
effectively will replace
those who don't.**

Q&A

