Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
RESTART

The Dark Side of the AI in 6G Networks

Marina Settembre

Fondazione Ugo Bordoni

RESTART

Plenary Dissemination Workshop

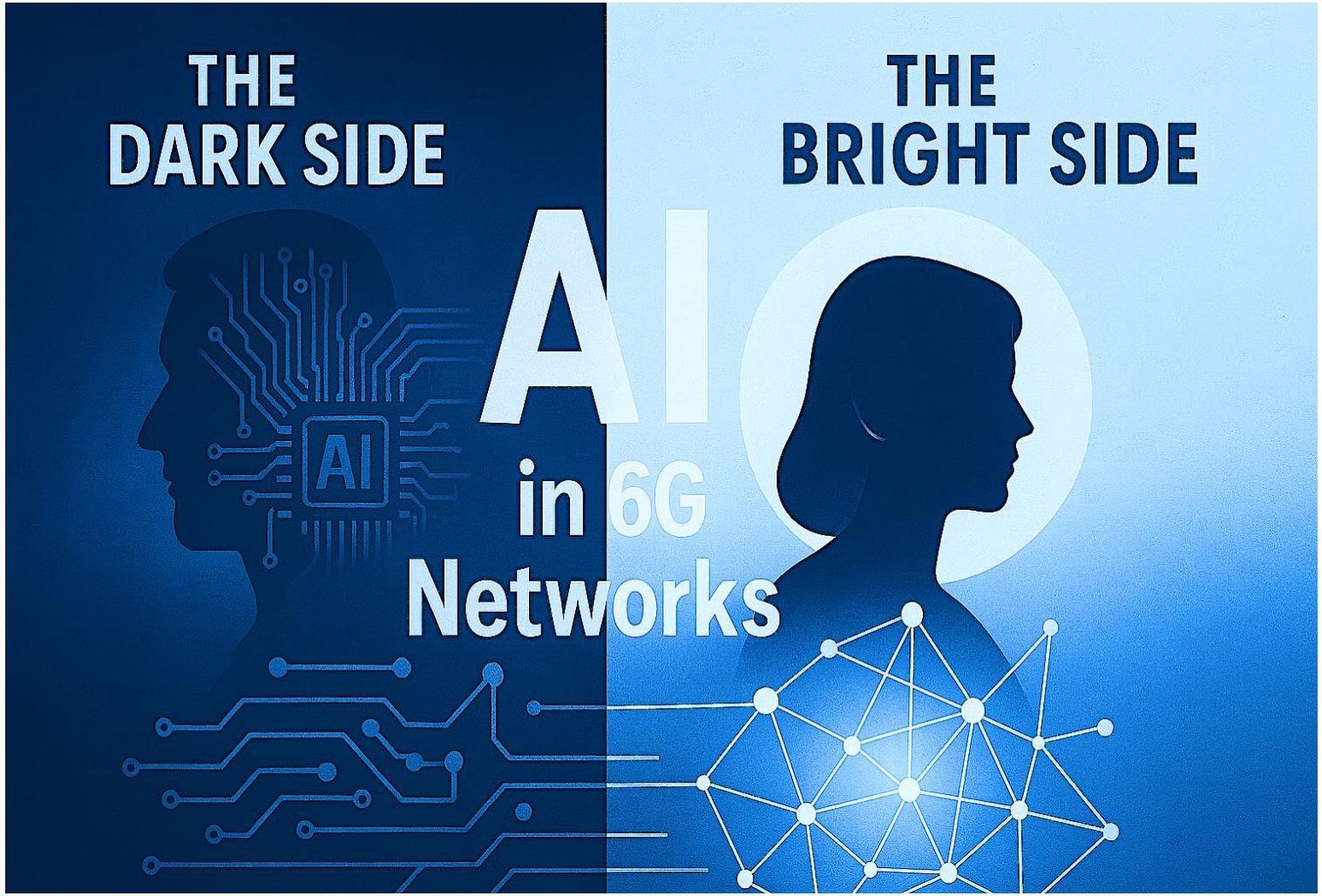
Hands-On with the Future of Telecommunications

June 30- July 2, 2025

Università degli Studi di Napoli Federico II

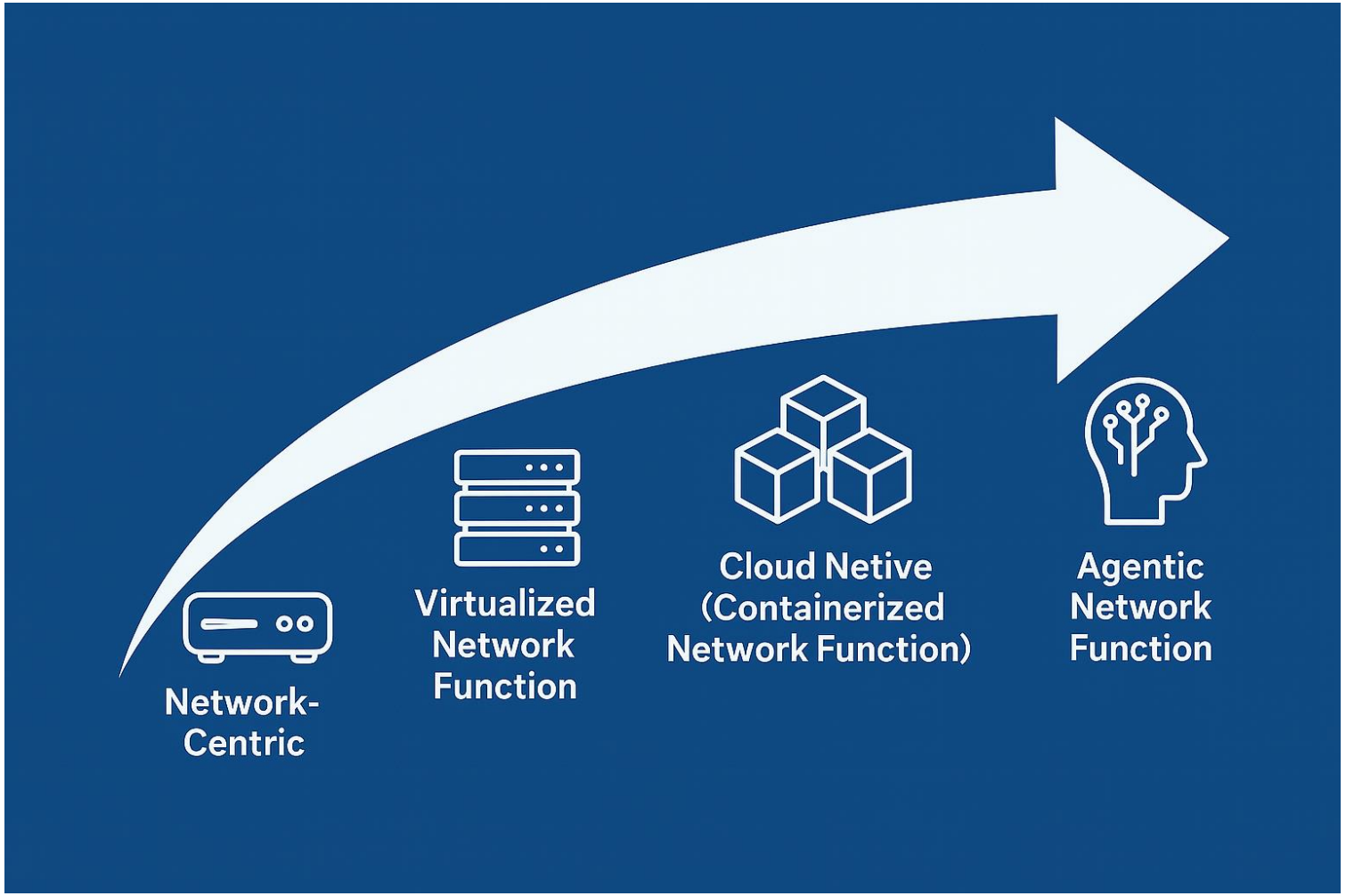
The Dark Side of AI

where the human being must shine brighter…



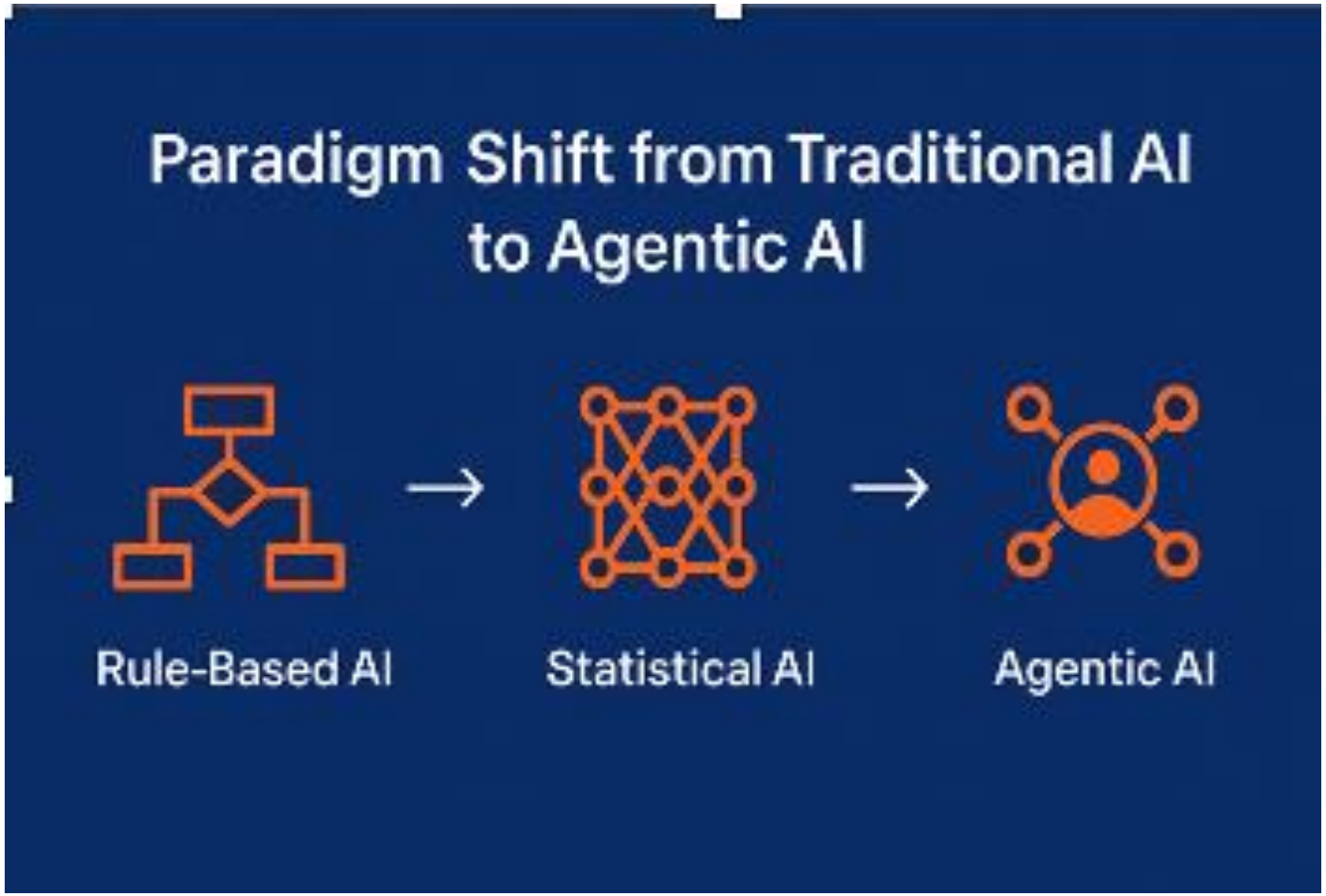
Note: All images have been generated with DALL·E – OpenAI (2025) by using concepts and prompts developed via ChatGPT by marina settembre

The Rise of Agentic AI Paradigm in TLC Networks



Risks Beyond Paradigm Shift

- In **tasks** the AI silently reprioritizes.
- In **business processes** it bypasses.
- In **KPIs** it deems "non-critical."
- In **policy rules** it doesn't know how to interpret.



The Invisible Algorithm

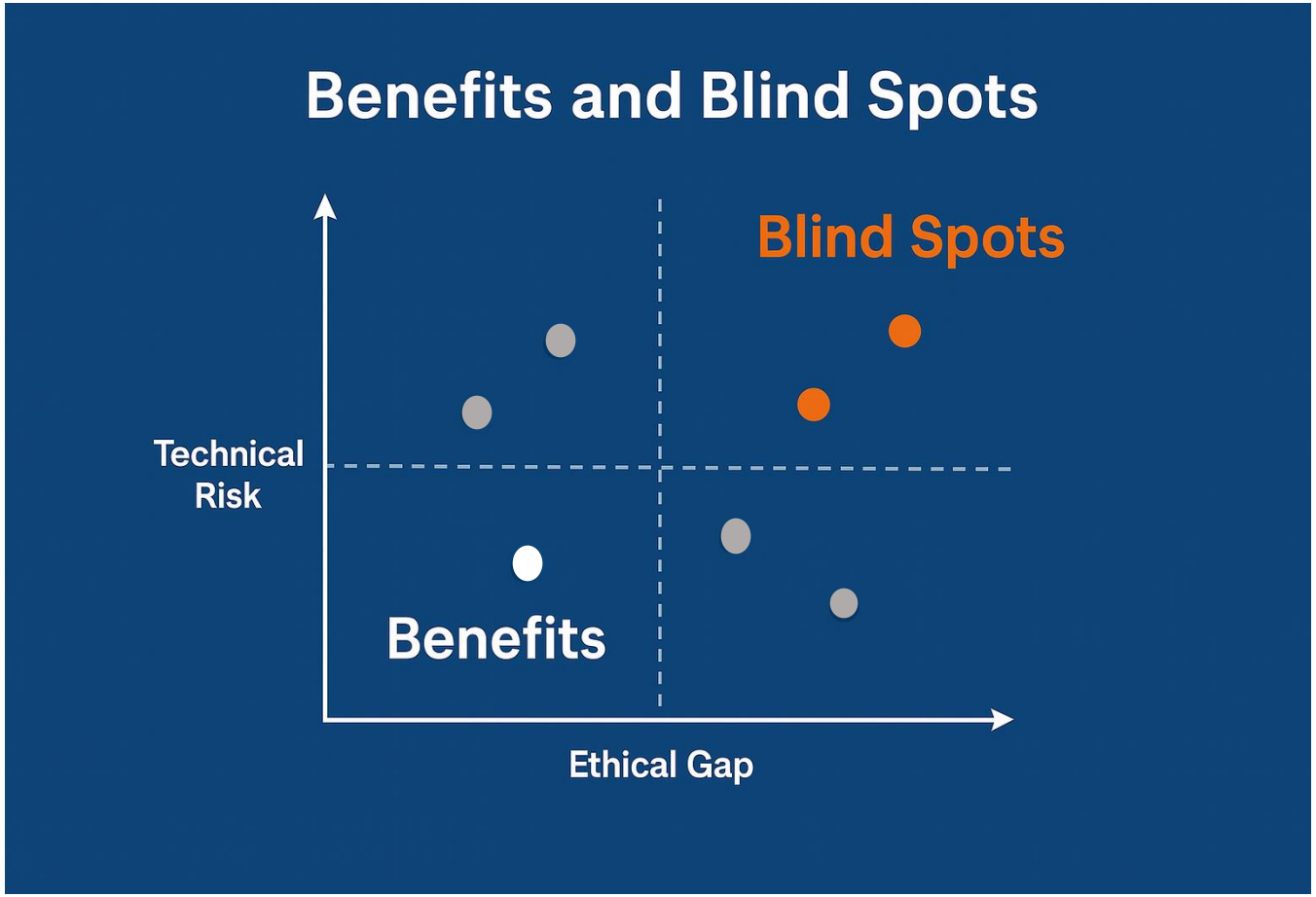
Who decides who gets connected or served ?

A silent algorithm or a human value system?

Benefits and Blind Spots with AI

Benefits Zone: AI delivers clear value, low risk, high efficiency

Blind spot: technically powerful, ethically fragile.



Multilevel Risk Chain

A systemic approach from Technical up to Fundamental Rights

TECHNICAL LAYER
- Vulnerability in MLmodel or configuration
- Data poisoning / telemetry manipulation
- Unexpected input (e.g. adversarial attacks

FUNCTIONAL LAYER (AI TASK EXECUTION)
- AI reprioritizes or misclassifies a critical network function
- Agent misalignment with operational goals
- Silent failure in decision logic (no allert triggered)

ORGANIZATIONAL LAYER
- Bypass of existing business rules or human in the loop control
- Misalignment with operational KPIs
- Automation overrides domain expert judgment

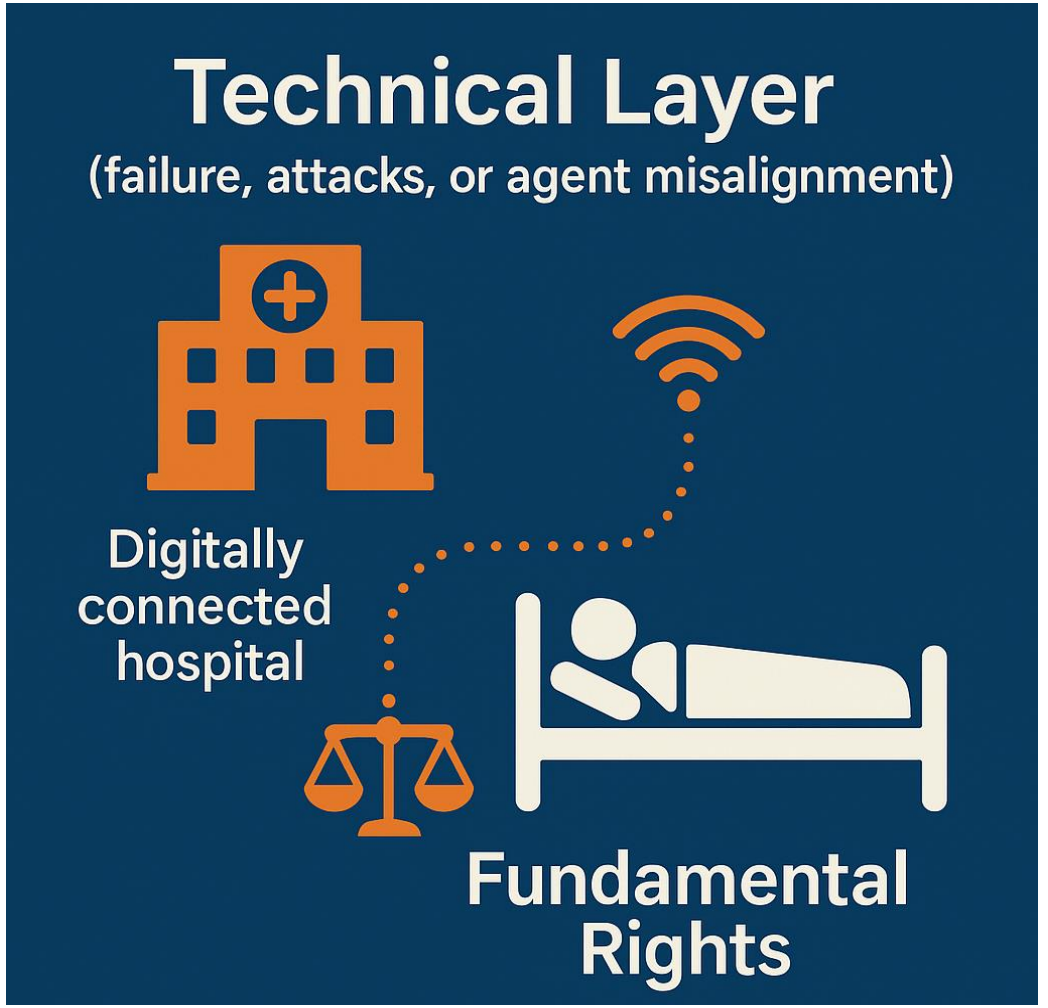
HUMAN IMPACT LAYER
- Indirect harm to people (e.g. outages of emergency services)
- Discriminatori routing / prioritization

Governance by Design

It is a call for new research directions to model harm before it happens, not only regulations

An Example from PNRR SERICS Spoke 7 - Eraclito Project

A preliminary cross-disciplinary exploration of how technical risks cascade into human consequences through a semantically rich ontology and a real-world test case.



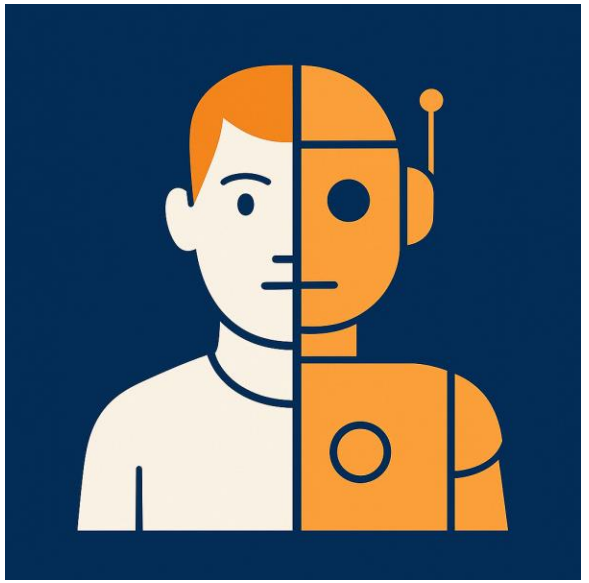
Closing Thoughts

In *The Cloven Viscount*, Calvino leaves us a warning for our times:

Let's stay whole.

…that also means

- *Lets'find a balance between overregulation and innovation*
- *Let's find a balance between leveraging AI and mitigating risks*
- *Let's connect the dots between disciplines, research programs, and across the ecosystem for a collective intelligence*



The challenge isn't to build perfect machines
it's to take responsibility for what machines decide.

RESTART