

AI-enabled Software Engineering, Software and Society

Assoc. Prof. Emil Alégroth

2026-10-23

SERL Sweden
Leading Software Engineering



VINNOVA

What is AI?

- **ML/AI**
 - Data
- **GenAI**
 - Prompts
- **Agent-workflows**
 - Agents
- **Agentic systems**
 - Learning



What are (some) use cases of GenAI?

- **Software artefact generation**
 - Requirements, code, tests
- **Document generation**
 - Code comments, synthesis, presentations
- **Information and knowledge**
 - Search, rubber duck, ideation
- **Management**
 - Synthesis, coordination, alignment



What are the common denominators?

- Artefact review
 - Shift right
- Automation
 - Agentification

Agent-builder frameworks (small sample)



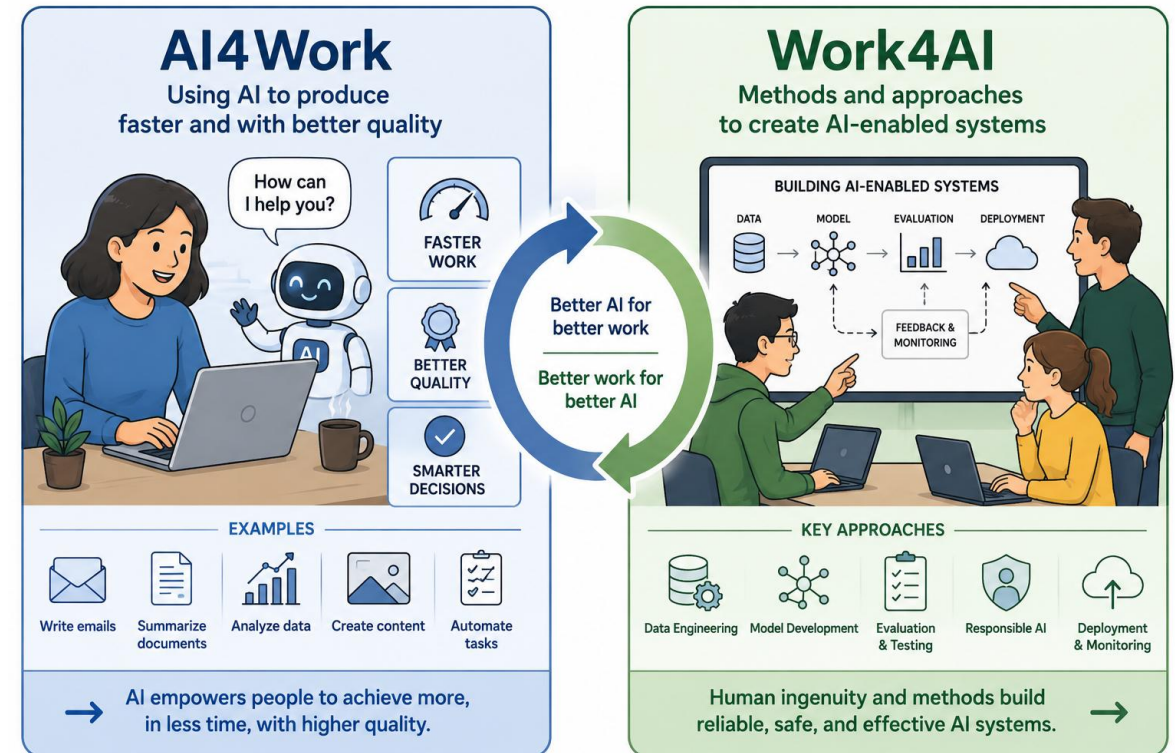
Properties of GenAI we **MUST** consider!

- Statistical model
- Non-deterministic
- Non-understanding
- Sycophantic
- Can be incorrect
- Limited memory
- Lacks agency



GenAI-enabled Software Engineering

- **AI4SE** – Tools, frameworks and methods that support development of software.
- **SE4AI** – Process and methods for developing GenAI-enabled software



Beyond technology

- **Economics** – Cost vs Value of GenAI?
- **Psychology** – Do we understand, accept and trust AI?
- **Process** – Human vs AI-agents? New roles?
- **Data** – Correct? Valid? "Good enough"?
- **Legal** – Responsibility? Compliance?
- **User experience** – Text vs speech vs background process?
- **Security** – Control security vs dynamic adaptation?



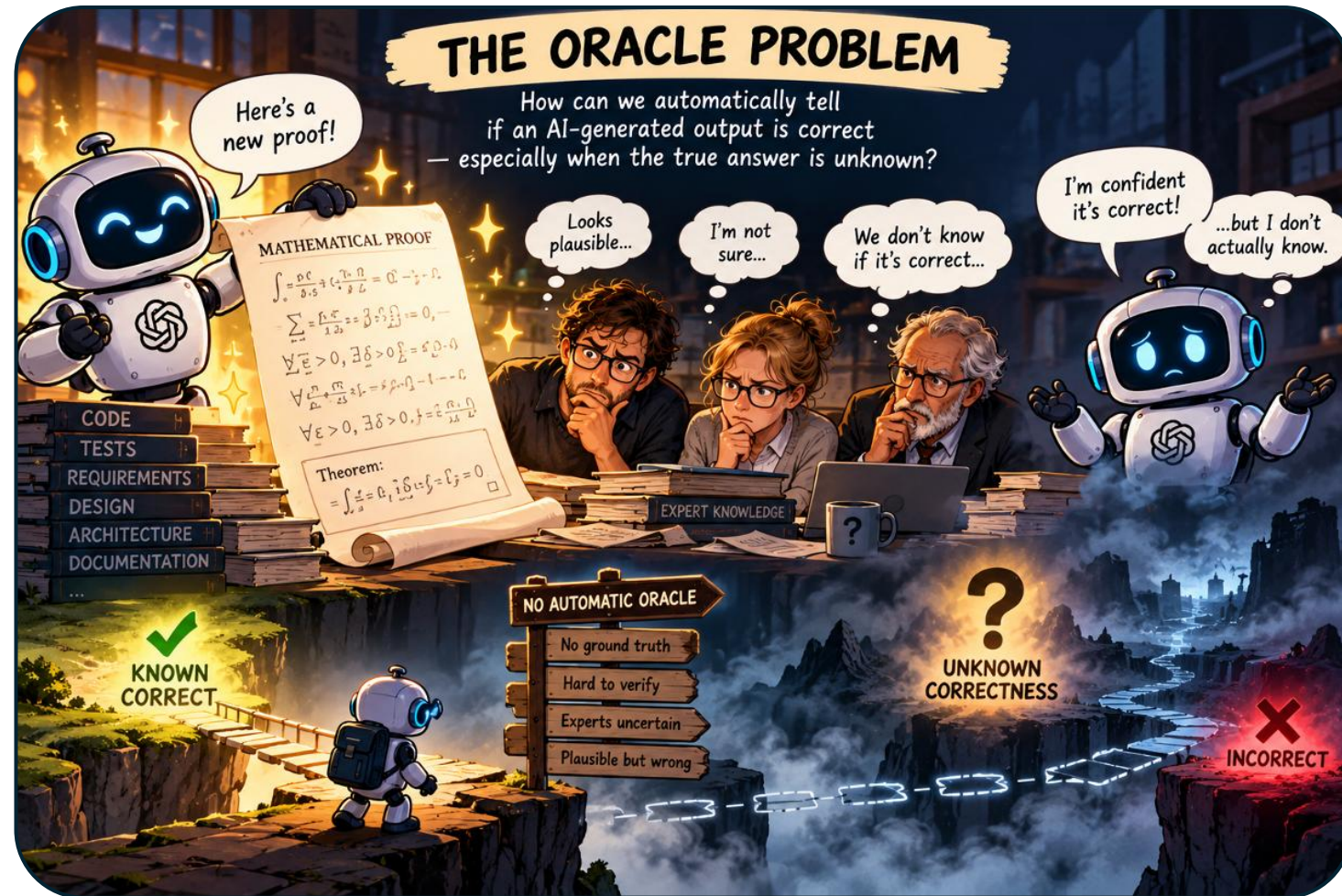
What's the fundamental problem?

- **The Oracle**

- How do we determine that something is correct?

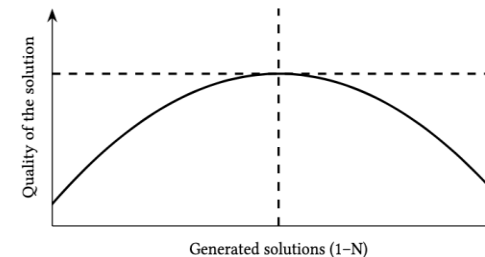
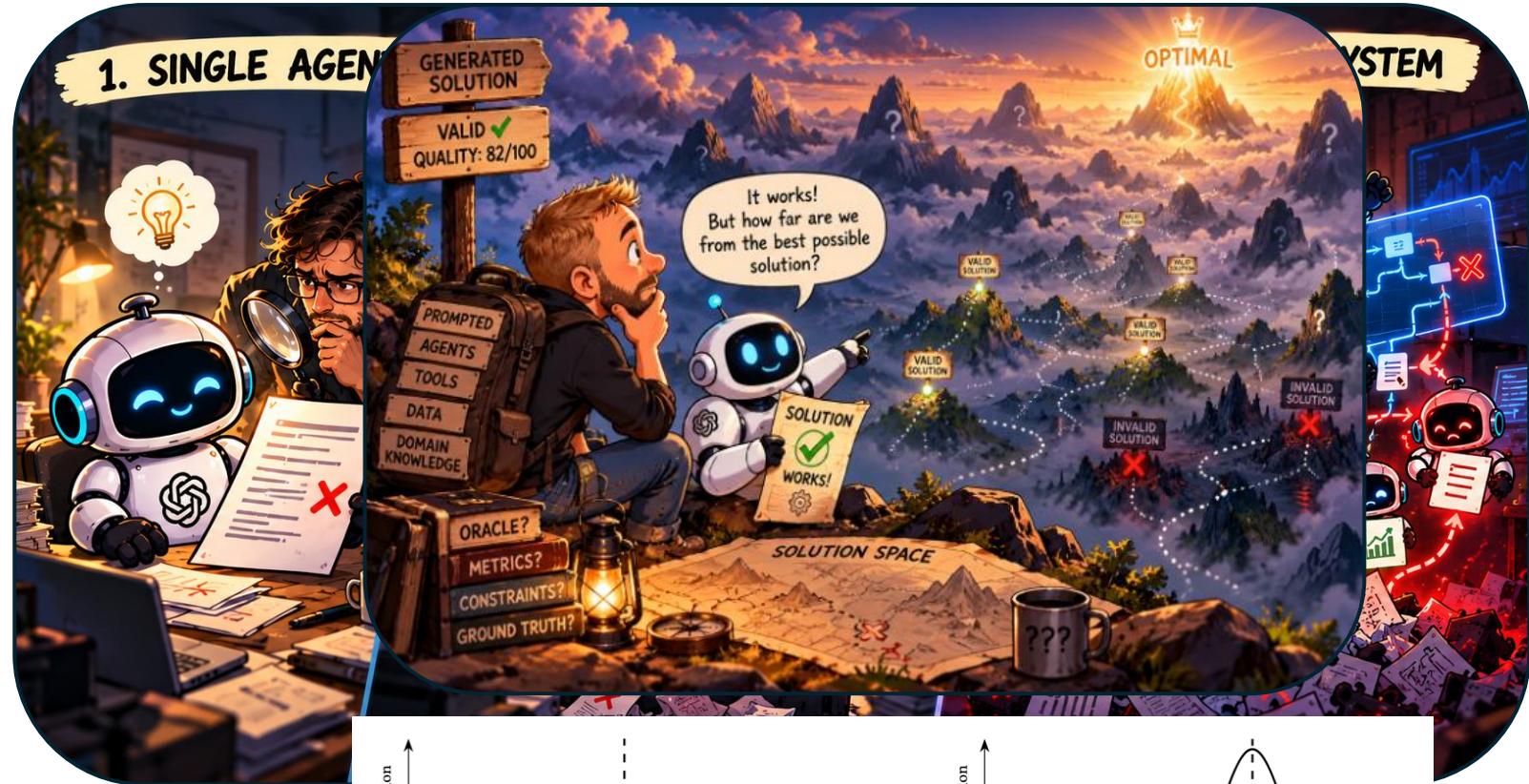
- **Example Oracle technologies**

- Semantic similarity
- LLM-as-a-judge
- Metamorphic oracles
- Differential oracles
- Statistical oracles
- Property-based oracles
- Probabilistic evaluation

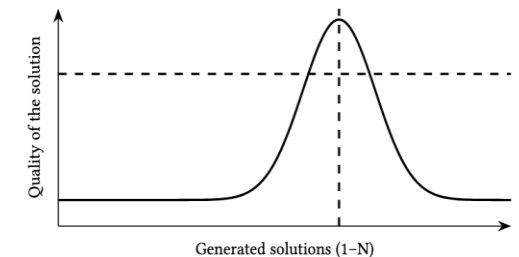


Compounded Oracle problems

- Agent
- AI workflow
- Agentic AI
- **Result correctness metrics?**
 - Distance



(a) Single-agent: broad applicability.



(b) Multi-agent: narrower scope, higher peak.

Short-term gains, long-time loss?

- **Linear efficiency gains**
 - Exponential quality degradation
- **Cognitive atrophy**
 - Less innovation
- **Workload balance**
 - More production in less time
- **Turnover**
 - Next generation senior developers
- **Cost**
 - Current ROI of GenAI globally



Additional problems

- **GenAI is transformative, not obviating!**
 - Industry is not transforming fast enough
- **Security and legal concerns**
 - Claude Fable and Mythos security
 - Alibaba AI “escaped” to mine bitcoin
 - OpenClaw’s ClawHub infected with worm
 - Lawyer in NY used AI references (May 20th)
 - OpenAI agents escaping their sandbox to cheat
- **What is true and what is hype?**



If AI is so problematic, why use AI?

- “You will not be replaced by AI, you will be outcompeted by someone who uses it”



If GenAI R&D ended today, we still have 10 years of utilization research to do!



GenAI in society

- AI positive OR AI negative
- AI is an enabler
- AI-enabled developer **vs** AI-enabled engineer



How is BTH solving it?

- T.A.R.G.E.T. (Vinnova)
 - Innovating a tool called Scout for GenAI-Enabled GUI testing (AI4SE)
- GENESIS (KKS)
 - Identifying best practices for generating code and documentation with GenAI (SE4AI)
- ThreMolia (Vinnova)
 - Integrating threat modeling with AI4SE
- S.E.R.T. (KKS)
 - Next generation Software Engineering practices (SE4AI/AI4SE)
- SESAM (KKS)
 - Security testing with GenAI and Security Quality Assurance (SE4AI/AI4SE)
- SERAI (KKS) – New project!
 - Verification and Validation of GenAI
- PRAISE (KKS) – New project!
 - Industrial school and research into GenAI4Education

Almost 100MSEK in research!

...but it's *just the start!*

SERL Sweden
Leading Software Engineering

Synteda GROUP PROACT

Spotify



FORTNOX
Småföretagens bästa vän



CYBER RANGES

Swedbank

Handelsbanken



TRAFIKVERKET

SONY

ERICSSON

itestra
be excellent

QESTIT



telenor



HERENCO



SAAB

Reinsight

SIEMENS

tolpagorni



AXIS
COMMUNICATIONS



FMV

TimePeopleGroup

Telia

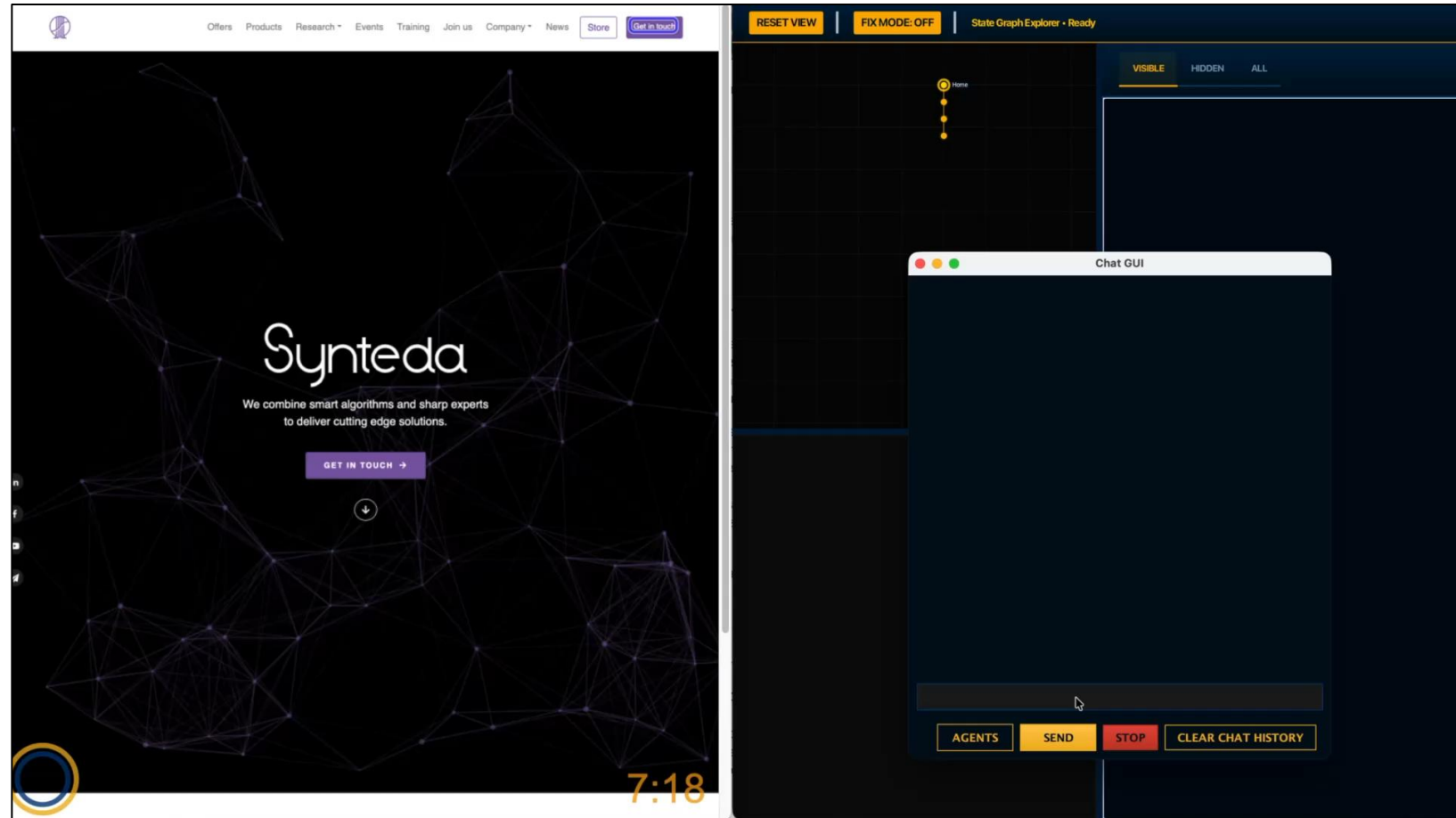


MaxKompetens



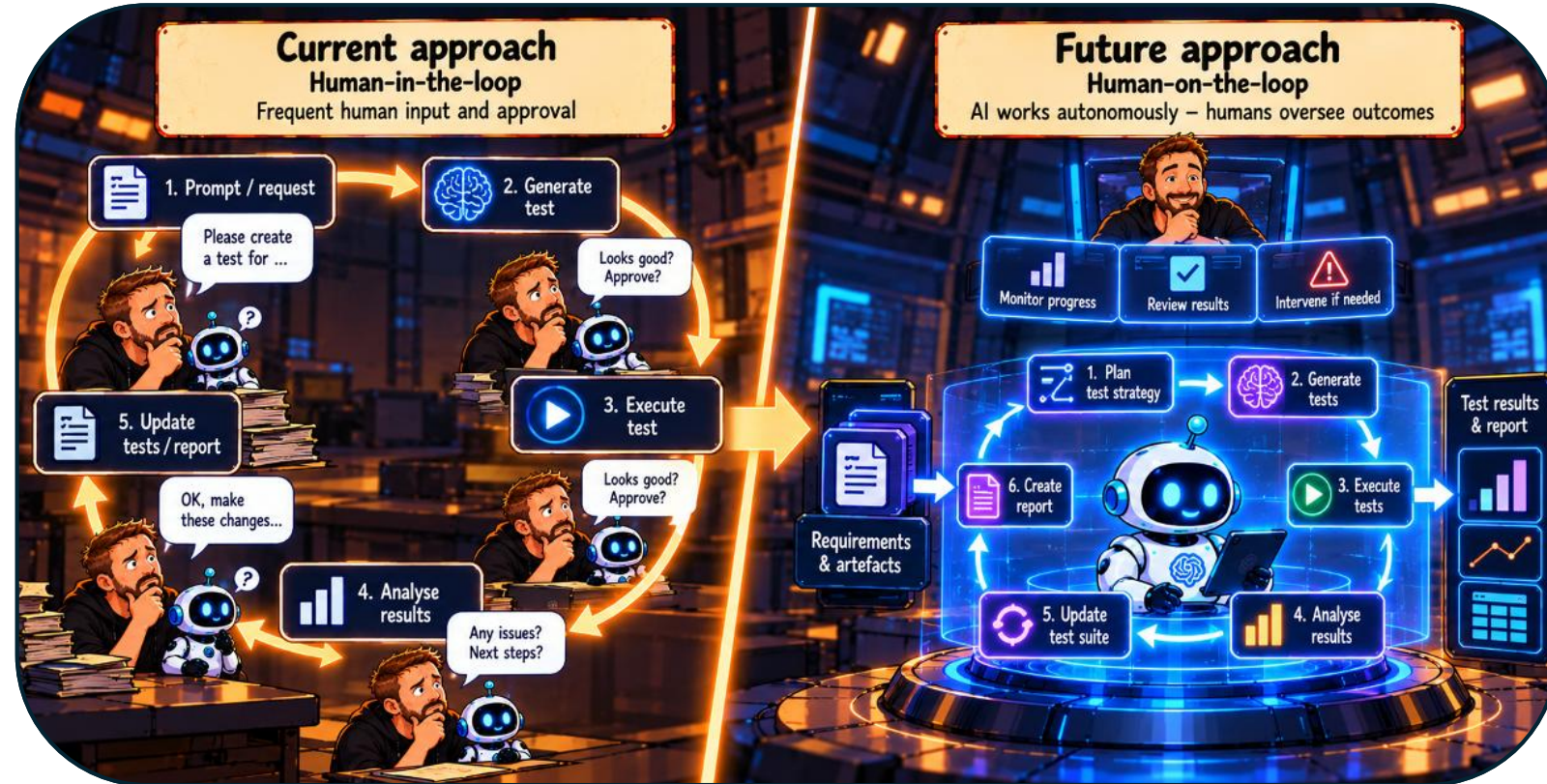
What is Scout/HiveMind?

- Record and replay
- What if we could generate the tests?
- **AI-Scout**
 - Multi-agent
 - Multi-modal
 - Multi-model



What's the vision?

- GenAI bottom-up automation
 - Agentic system
- Human-ON-the-loop
 - From review to oversee
 - Shift-right



Summary

Properties of GenAI we **MUST** consider!

- Statistical model
- Non-deterministic
- Non-understanding
- Sycophantic
- Can be incorrect
- Limited memory
- Lacks agency



What's the fundamental problem?

- **The Oracle**
 - How do we determine that something is correct?
- **Example Oracle technologies**
 - Semantic similarity
 - LLM-as-a-judge
 - Metamorphic oracles
 - Differential oracles
 - Statistical oracles
 - Property-based oracles
 - Probabilistic evaluation



If AI is so problematic, why use AI?

- “You will not be replaced by AI, you will be outcompeted by someone who uses it”



How is BTH solving it?

- **T.A.R.G.E.T. (Vinnova)**
 - Innovating a tool called Scout for GenAI-Enabled GUI testing (AI4SE)
- **GENESIS (KKS)**
 - Identifying best practices for generating code and documentation with GenAI (SE4AI)
- **ThreMolia (Vinnova)**
 - Security and threat modelling of LLM-integrated applications (SE4AI/AI4SE)
- **S.E.R.T. (KKS)**
 - Next generation Software Engineering practices (SE4AI/AI4SE)
- **SESAM (KKS)**
 - Security testing with GenAI and Security Quality Assurance (SE4AI/AI4SE)
- **SERAI (KKS) – New project!**
 - Verification and Validation of GenAI
- **PRAISE (KKS) – New project!**
 - Industrial school and research into GenAI4Education



Questions?

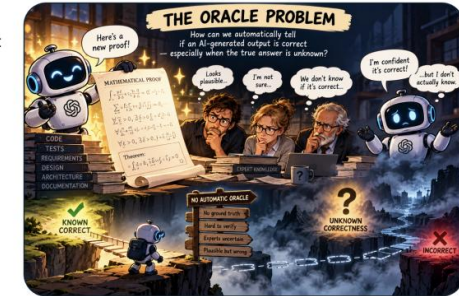
Properties of GenAI we **MUST** consider!

- Statistical model
- Non-deterministic
- Non-understanding
- Sycophantic
- Can be incorrect
- Limited memory
- Lacks agency



What's the fundamental problem?

- **The Oracle**
 - How do we determine that something is correct?
- **Example Oracle technologies**
 - Semantic similarity
 - LLM-as-a-judge
 - Metamorphic oracles
 - Differential oracles
 - Statistical oracles
 - Property-based oracles
 - Probabilistic evaluation



Contact

Emil.Alegroth@BTH.se

If AI is so problematic, why use AI?

- “You will not be replaced by AI, you will be outcompeted by someone who uses it”



How is BTH solving it?

- **T.A.R.G.E.T. (Vinnova)**
 - Innovating a tool called Scout for GenAI-Enabled GUI testing (AI4SE)
- **GENESIS (KKS)**
 - Identifying best practices for generating code and documentation with GenAI (SE4AI)
- **ThreMolia (Vinnova)**
 - Security and threat modelling of LLM-integrated applications (SE4AI/AI4SE)
- **S.E.R.T. (KKS)**
 - Next generation Software Engineering practices (SE4AI/AI4SE)
- **SESAM (KKS)**
 - Security testing with GenAI and Security Quality Assurance (SE4AI/AI4SE)
- **SERAI (KKS) – New project!**
 - Verification and Validation of GenAI
- **PRAISE (KKS) – New project!**
 - Industrial school and research into GenAI4Education



VINNOVA