

Agentic Software Engineering Will Eat the World

AI-based Systems as the new OS of Society

Keynote @ Agentic Eng. Workshop 2026, Rio de Janeiro, Brasil

Robert Feldt, robert.feldt@chalmers.se

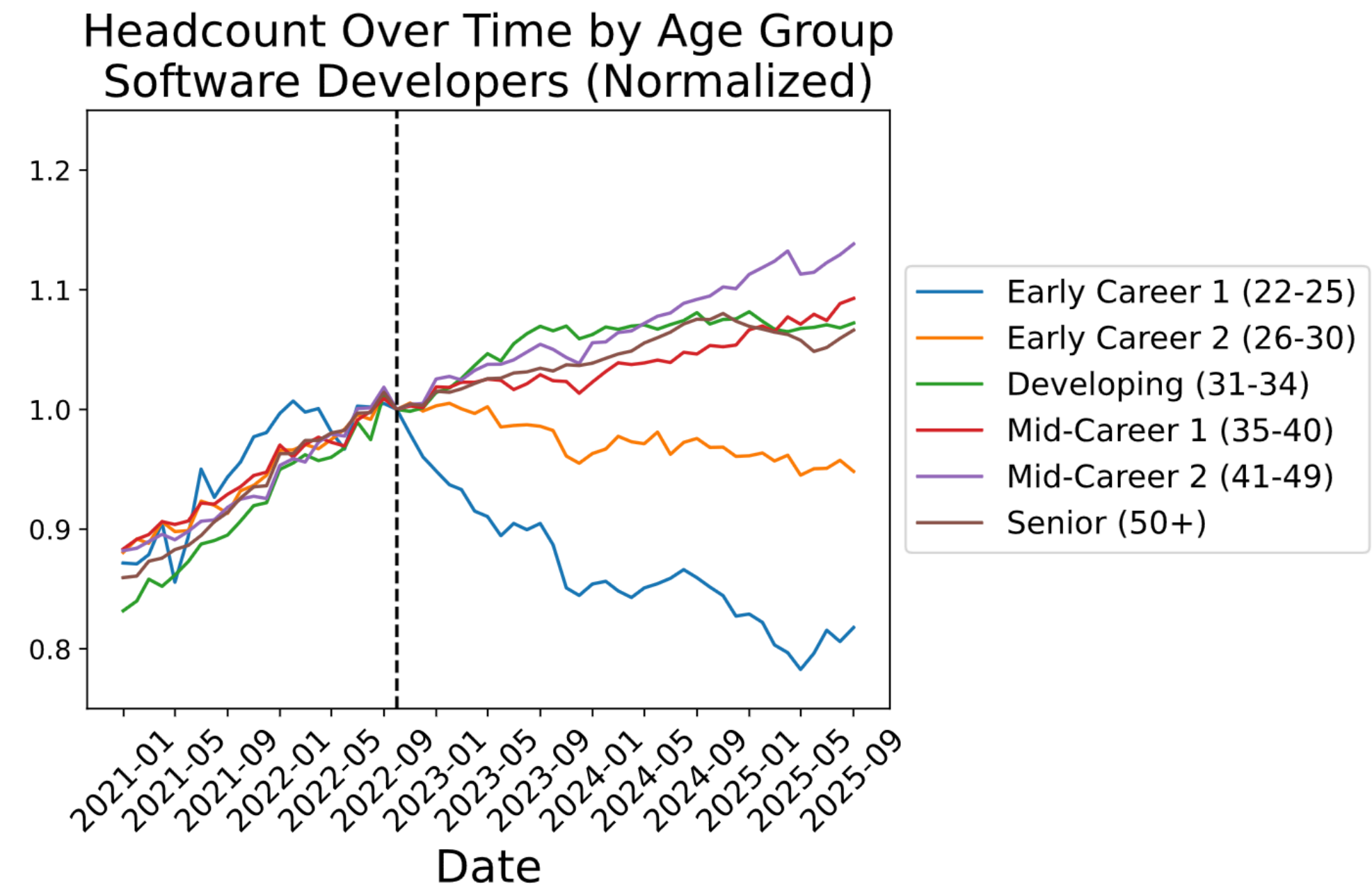
Chalmers University of Tech, Gothenburg, Sweden

The shrinking story

- Coding agents become more capable
- (Junior) roles contract, thin out, disappear
- End-user programming seems ready to bypass experts
- Senior developers disillusioned or depressed
- AI4SE & Agentic systems increasingly automates core SE activities

“Experienced open-source developers using early-2025 AI tools took 19% longer on average in a randomized trial”

- METR, “Measuring the Impact of Early-2025 AI...”, July 2025



[Brynjolfsson et al, “Canaries in the Coal Mine?”]

Industrial Multi-Agent LLMs

Automating a Complete Software Test Process Using LLMs: An Automotive Case Study

Shuai Wang¹, Yinan Yu¹, Robert Feldt¹, Dhasarathy Parthasarathy²

¹ Chalmers University of Technology ² Volvo Group

Gothenburg, Sweden

shuaiwa@chalmers.se, yinan@chalmers.se, robert.feldt@chalmers.se, dhasarathy.parthasarathy@volvo.com



Early agentic LLMs4testing

Towards Autonomous Testing Agents via Conversational Large Language Models

Robert Feldt

Chalmers University of Technology

robert.feldt@chalmers.se

Sungmin Kang

KAIST

sungmin.kang@kaist.ac.kr

Juyeon Yoon

KAIST

juyeon.yoon@kaist.ac.kr

Shin Yoo

KAIST

shin.yoo@kaist.ac.kr

GoNoGo: An Efficient LLM-based Multi-Agent System for Streamlining Automotive Software Release Decision-Making

Arsham Gholamzadeh Khoei^{1,2}[0000-0002-5130-5520], Yinan

Yu¹[0000-0002-3221-7517], Robert Feldt¹[0000-0002-5179-4205], Andris Freimanis²,

Patrick Andersson Rhodin², and Dhasarathy

Parthasarathy²[0000-0002-3620-8589]

Autonomous Large Language Model Agents Enabling Intent-Driven Mobile GUI Testing

Juyeon Yoon

KAIST

juyeon.yoon@kaist.ac.kr

Robert Feldt

Chalmers University of Technology

robert.feldt@chalmers.se

Shin Yoo

KAIST

shin.yoo@kaist.ac.kr

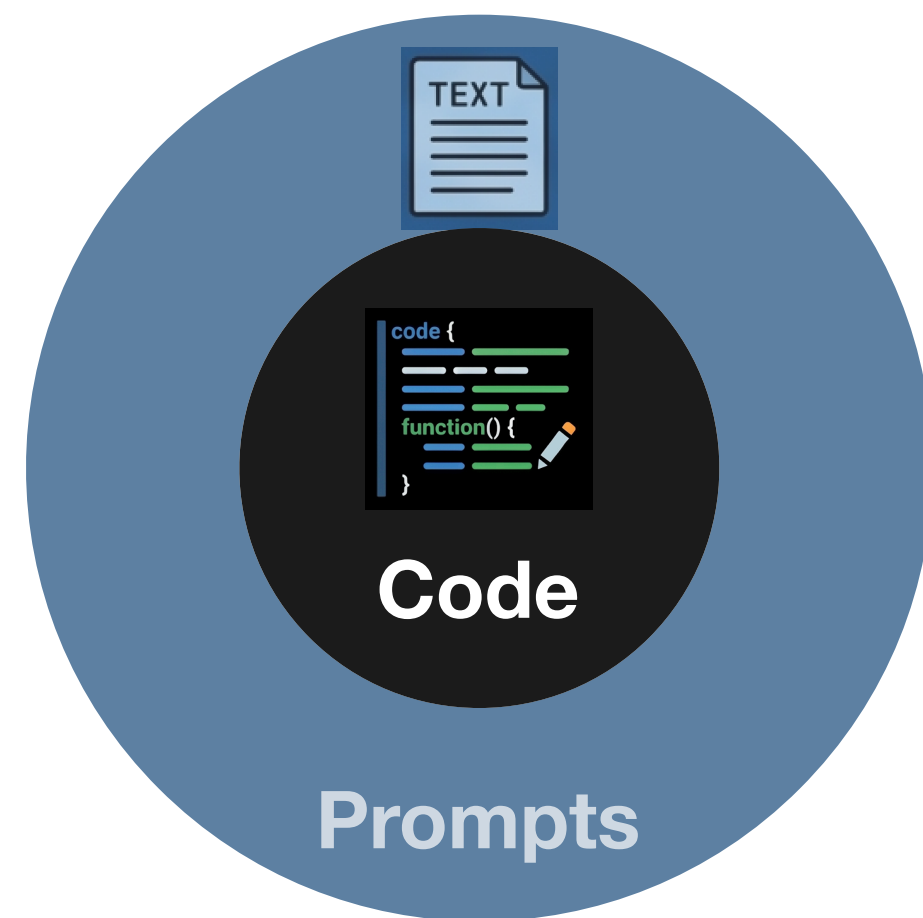
Ongoing:

- task-specific memory design (tied to companies' context eng. challenges)
- general code agents > custom MALLM (Volvo Infotainment GUI testing)
- organizational adaptation and strategy via skills

The expanding stack

“Firms are moving toward human-agent teams and redesigned work structures”

- MS Worklab, “Year of the Frontier Firm”, 2025

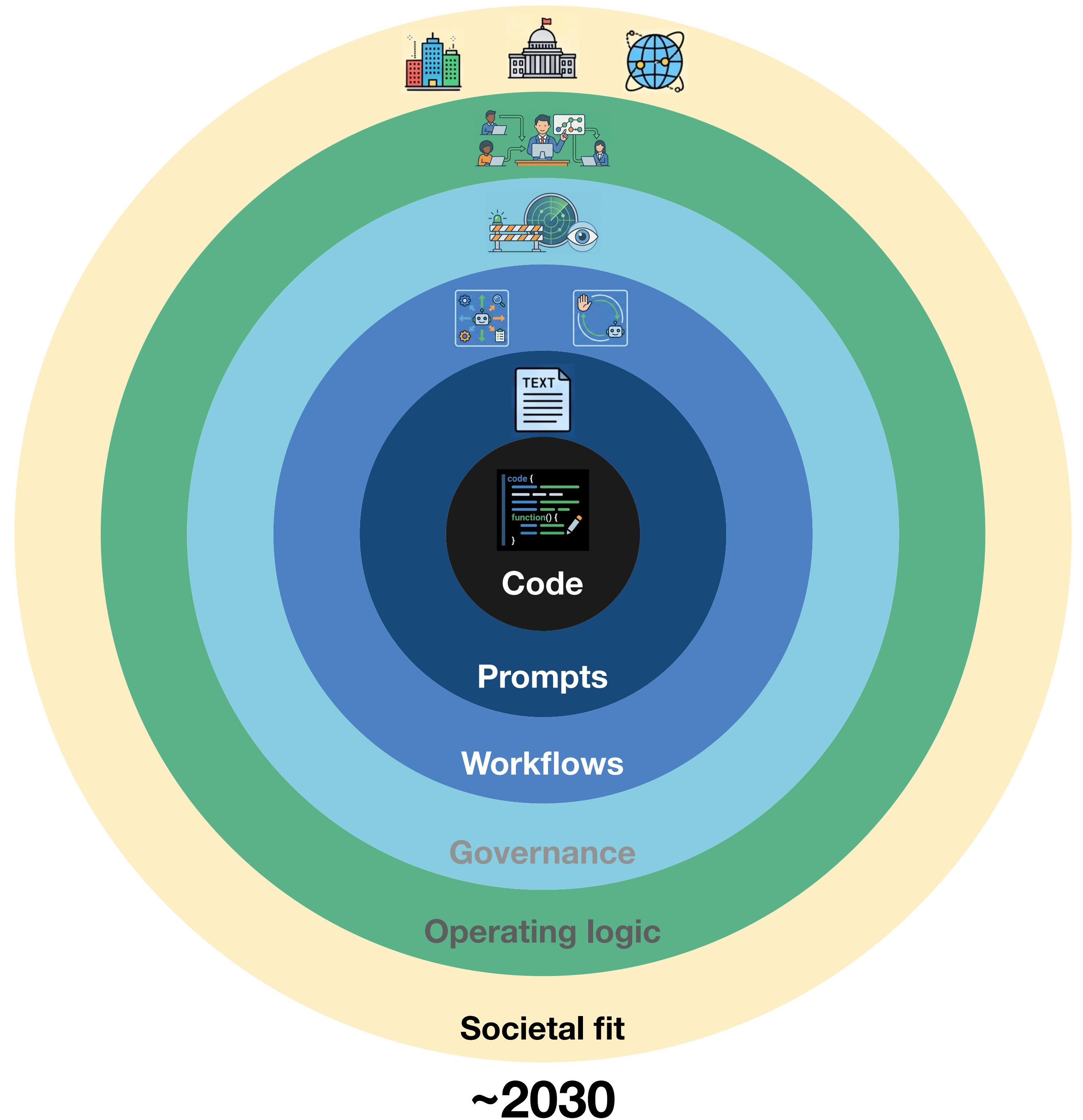


~2020



“AI behaves as an amplifier of existing organizational strengths and weaknesses”

- DORA, “2025 State of AI-Assisted SW Dev Report”

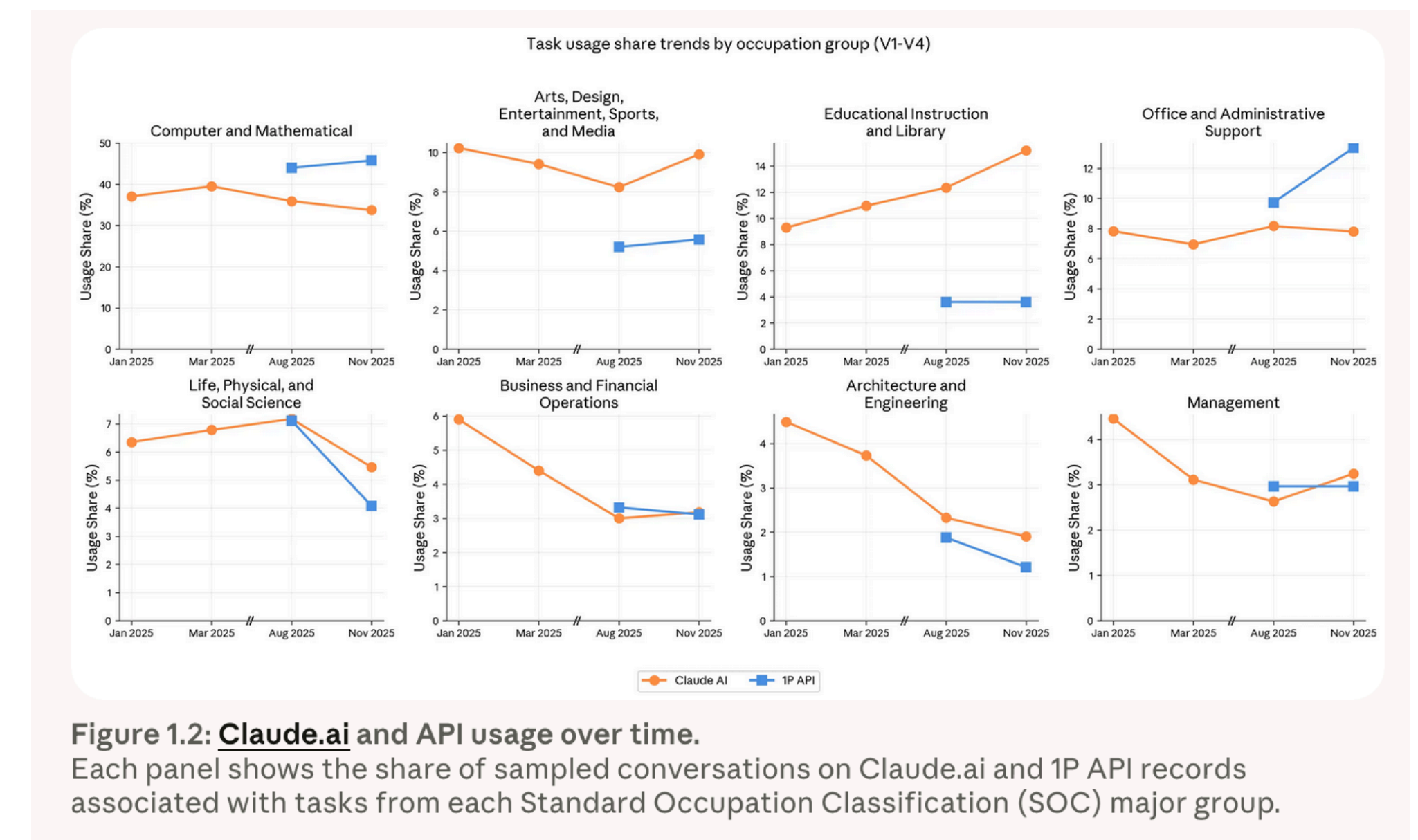
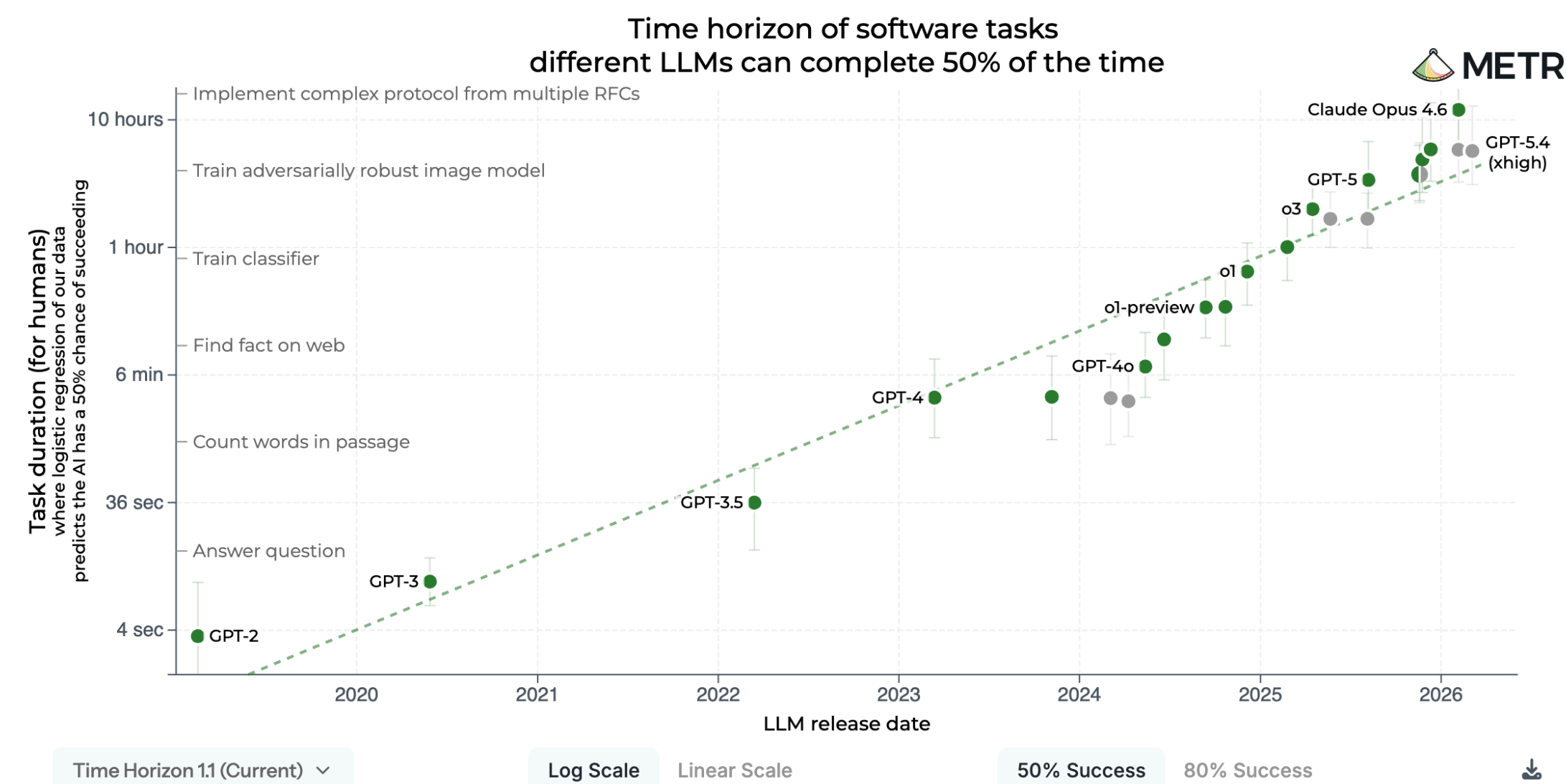


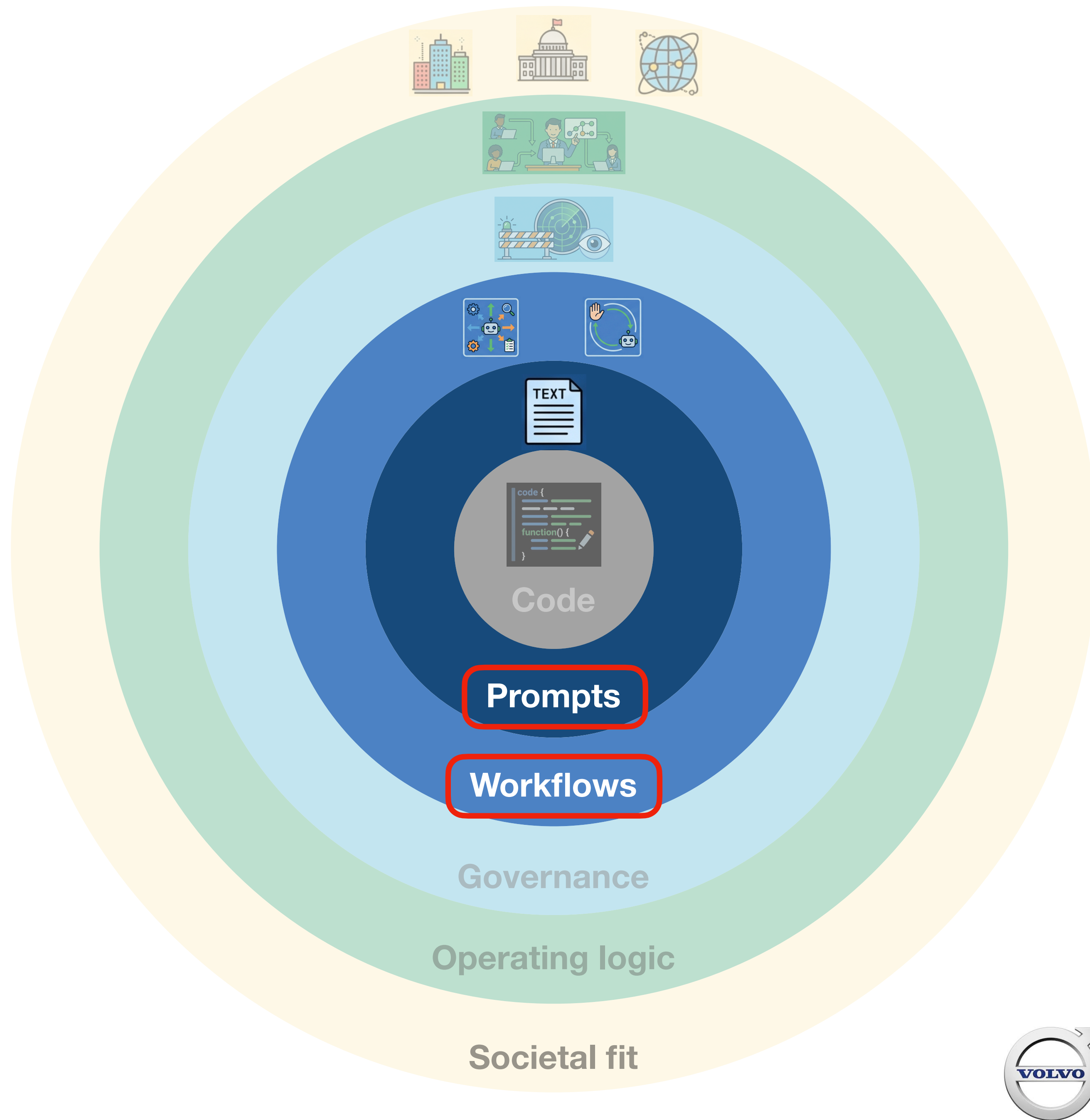
Societal fit

~2030

Imperfect agents already matter

- Capability trajectory is rising fast enough that static snapshots mislead
- Real-world usage already concentrates in SE-adjacent tasks
- Measurable effects in serious industrial settings
- But: realistic software work can still expose limits
- Still: Compounding availability beats isolated brilliance





Technical plumbing rings (2&3):

Prompts, NL specs, agent workflows, tool use, retrieval pipelines, multi-agent protocols, human-agent loops

Many concrete challenges, of course:

Context packaging, task-specific memory, custom vs general architectures, ...

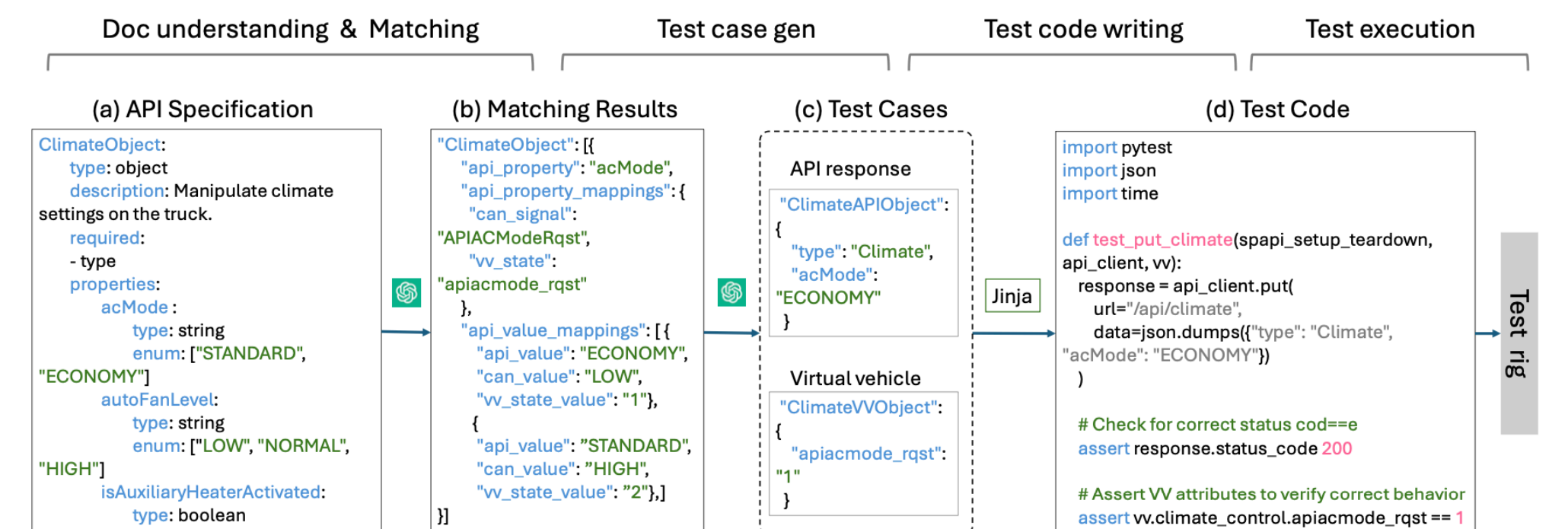
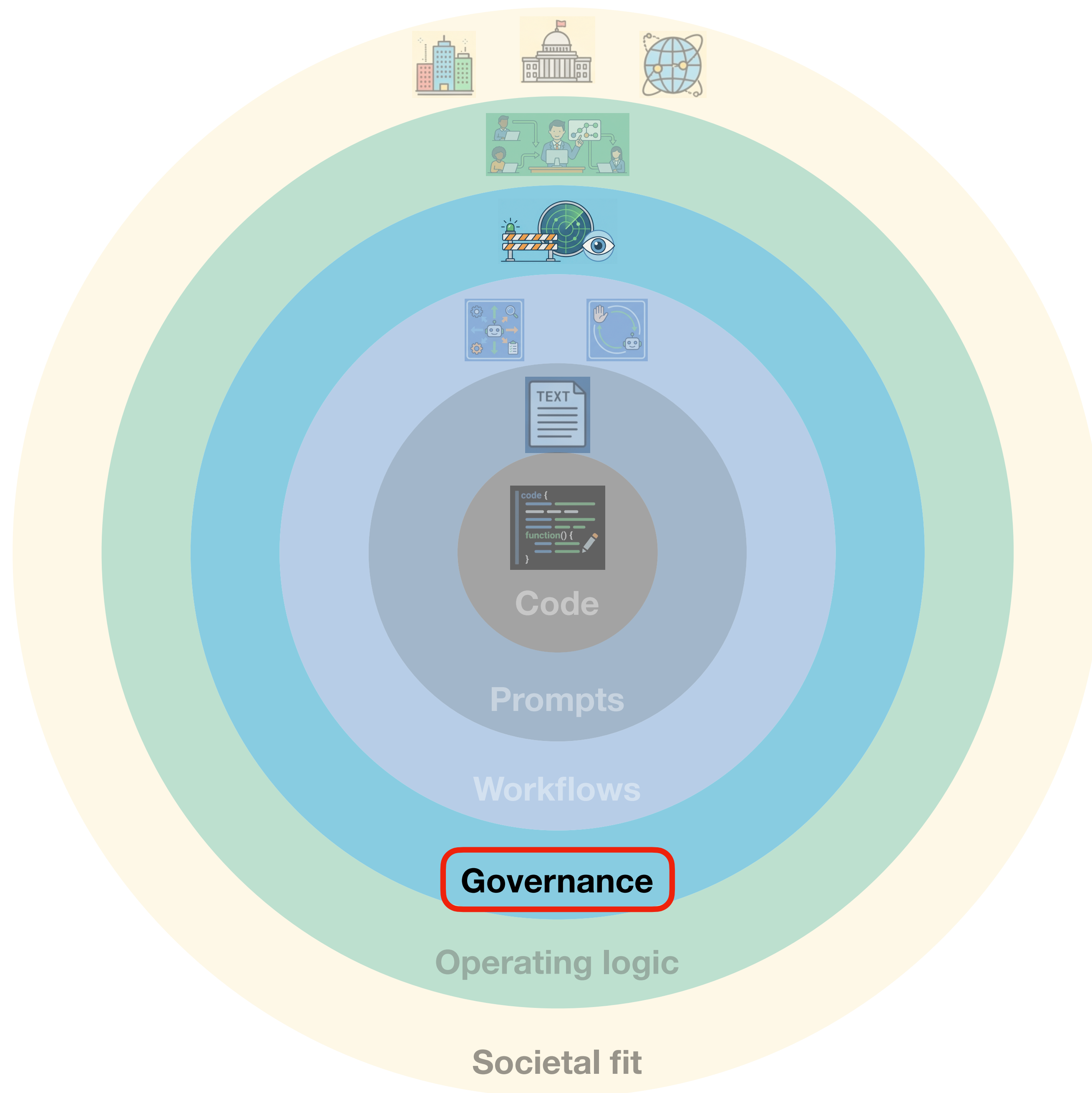


Fig. 5. Architecture and workflow of SPAPI-Tester: The pipeline largely preserves the manual process and selectively uses LLMs to automate discrete steps.



Control layer (4):

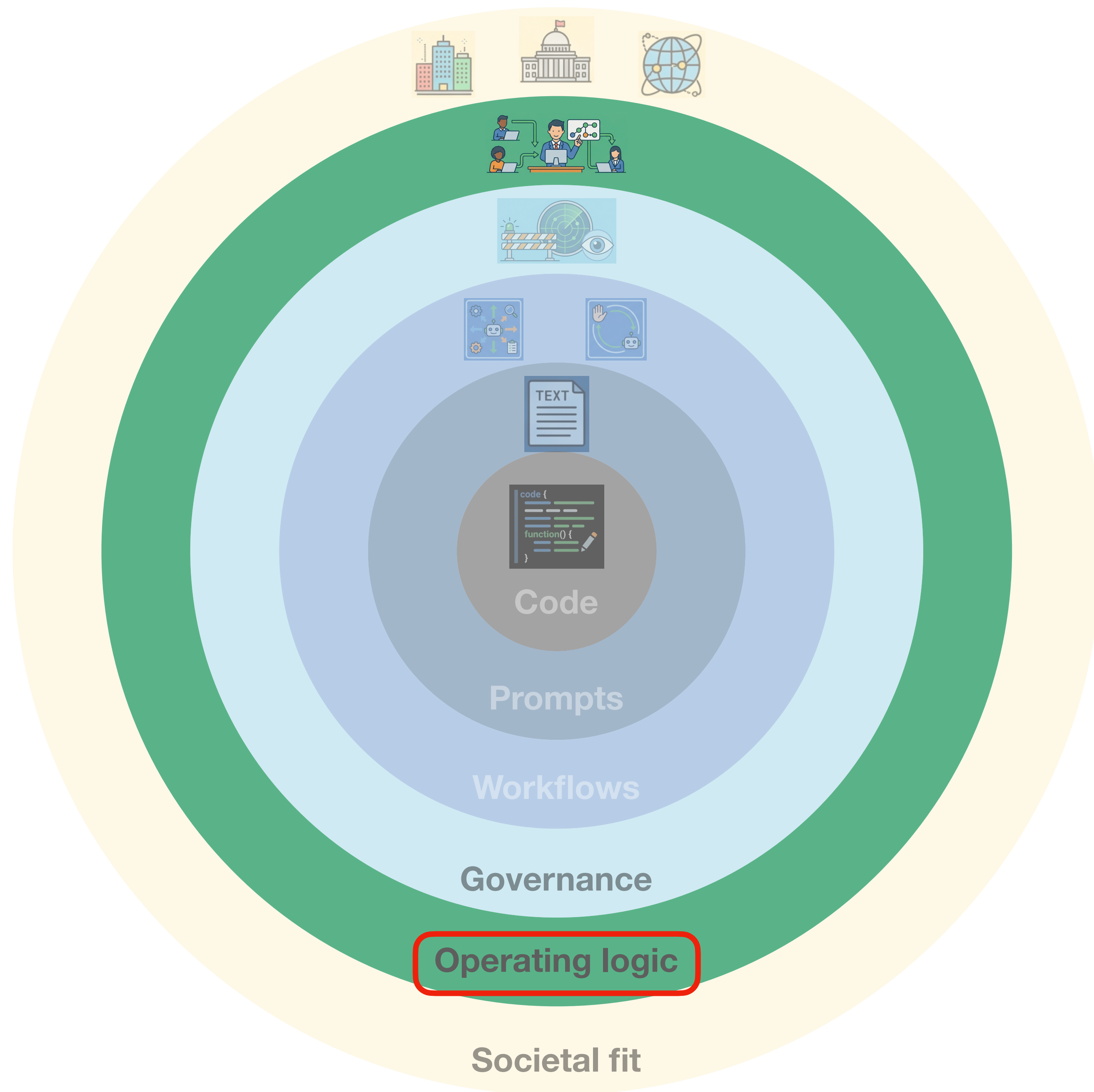
guardrails, monitoring, evaluation harnesses, human-in-the-loop design, failure-response protocols, policy layers

Workflows (3) is **how agents do work**,

Governance (4) is **how we control it works**

Governance (4) also connects to **management & strategy**

Artificial Intelligence Risk Management Framework (AI RMF 1.0)



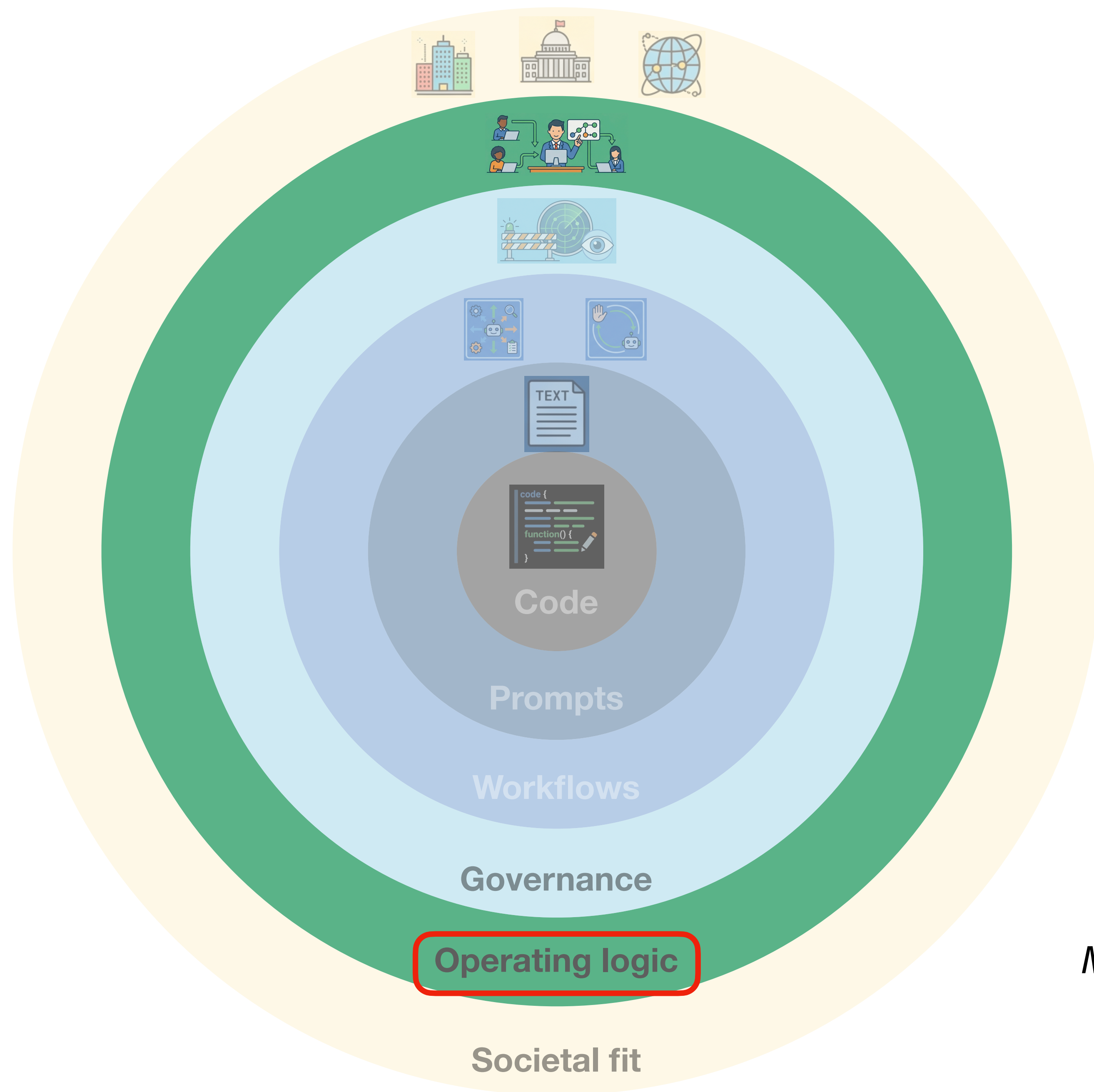
Operating logic (5):

Decision preparation, knowledge capture and reuse, coordination, value creation, company-wide learning and role exchange

The Cybernetic Teammate: A Field Experiment on Generative AI Reshaping Teamwork and Expertise

Fabrizio Dell'Acqua, Charles Ayoubi, Hila Lifshitz, Raffaella Sadun, Ethan Mollick, Lilach Mollick, Yi Han, Jeff Goldman, et al. (View all)

“Individuals using AI matched the performance of teams without AI”, field experiment of 776 employees on product innovation



Operating logic (5):

Routines, delegation, oversight, trust also come into play here and become engineering problems since part of the semi-executable organisational operating logic.

Real value here is not only in implementing and making current procedures more “efficient”

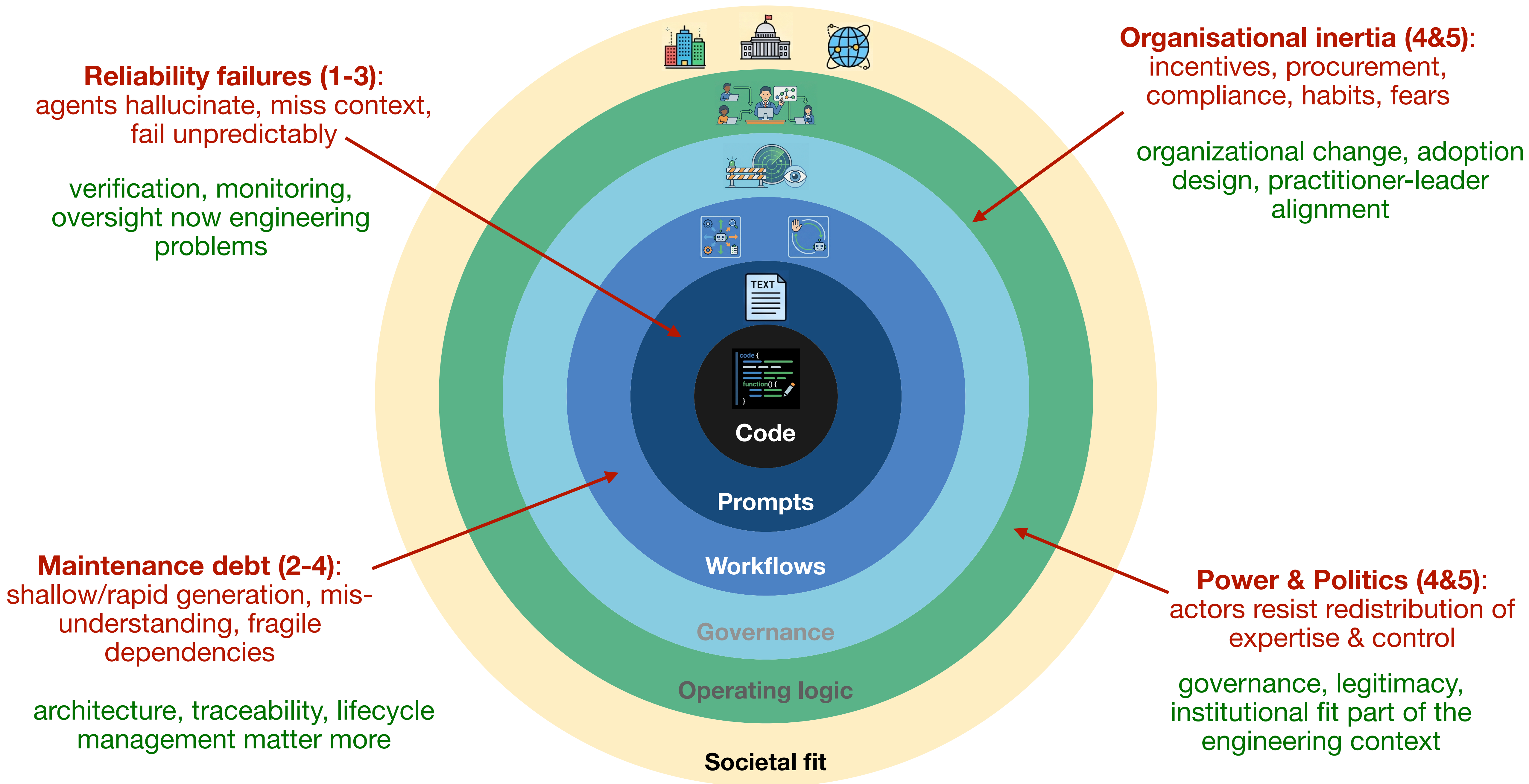
It’s in freeing up more creativity, experimentation, individualisation & fast learning

GoNoGo: An Efficient LLM-based Multi-Agent System for Streamlining Automotive Software Release Decision-Making

Arsham Gholamzadeh Khoei^{1,2}[0000–0002–5130–5520], Yinan Yu¹[0000–0002–3221–7517], Robert Feldt¹[0000–0002–5179–4205], Andris Freimanis², Patrick Andersson Rhodin², and Dhasarathy Parthasarathy²[0000–0002–3620–8589]

MALLM supporting release decision-making at Volvo

Freed up creativity and exploration.



Implications

Preserve	Purify
Explicit reasoning about requirements, constraints, assumptions	Practices optimized for manual code production
Modularity and interface discipline	Process artifacts for low-bandwidth human coordination
Validation and verification	Code-centric notions of quality that ignore prompts, policies, workflows
Traceability between intent, implementation, behavior	Strict phase boundaries in interactive human-agent loops
Lifecycle thinking and maintenance awareness	Narrow views of who counts as a system builder
Socio-technical realism about teams, communication, failure	Techno-focus and -optimism

Implications for groups

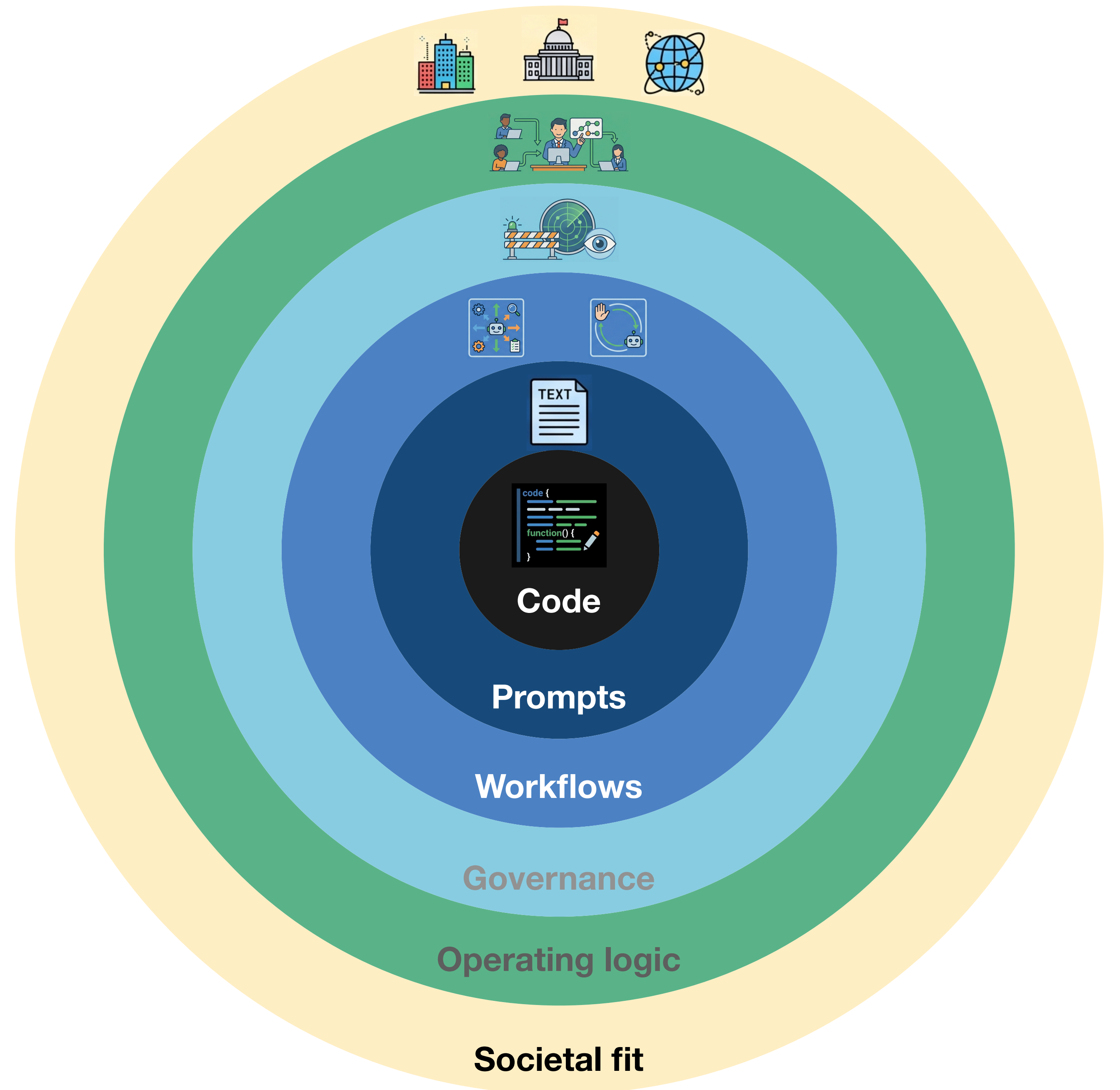
- Researchers:
 - Revisit whether your target is stable or a legacy bottleneck being automated away
 - Treat prompts, policies, workflows as first-class engineering objects
 - Ask whether parts of AI4SE are aimed at a shrinking conception of SE
- Practitioner (bottom-up)
 - Having engineering discipline becomes more important, not less
 - Scarce skill is deciding what to build, how to validate it, and how to integrate it into organizations that must trust and maintain it over time
- Leader (top-down):
 - AI is not a collection of efficiency tools
 - Question is whether your organization's operating logic is fit for this new world where knowledge work becomes partially encoded, executed, and improved
 - Waiting for certainty may itself be a losing strategy

Agentic software engineering will eat the world — not because code disappears, but because software-like, AI-based systems increasingly mediate how institutions operate

Public sector lags; full "OS of society" meaning only emerges when operating-logic changes scale across enough organizations and nations

Software engineering therefore becomes more central, while also needing to change

robert.feldt@chalmers.se



Backup/extra slides

Old as presented



New as in paper



