GN5-2 WP6 study: Intelligent Networks - The Rise of Generative AI in Network Management

Pavle Vuletić, UoB/AMRES
WP6 leader

16th SIG-NGN, KIT Karlsruhe 9-10.10.2025.

Motivation for WP6 GenAI study

Network management path from manual -> OSS/BSS -> automation -> ML -> GenAI

ChatGPT appeared in 2022, and became fastest-growing consumer software

New buzzword, there is a sense that “something should be done about this”

If it can write text or code, can it write network configurations, scripts?

Can it really help in other network management fields?

The use of AI in network management is the key topic of scientific conferences in the last two years, also in IRTF

IRTF NMRG Interim meeting on AI 8.10.2025.

Conference theme: theme “AI for Management and Management for AI”

WP6 AI-related activities

Chatbot for Network eAcademy (completed in GN5-1)

New incubator projects:

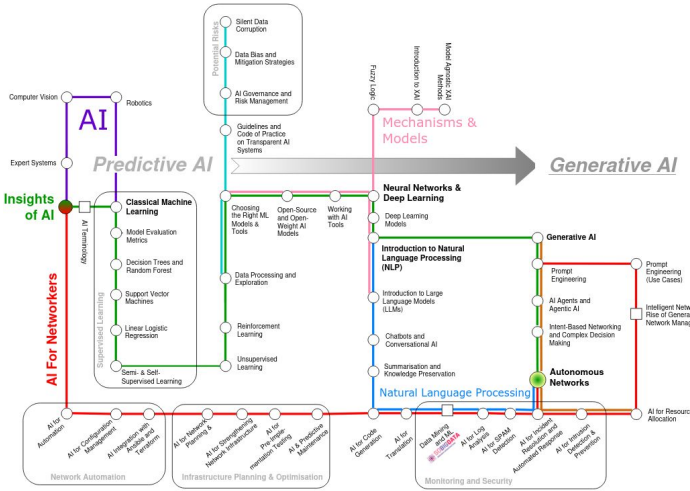
- LogLLM: Automated Log Analysis and Management based on LLMs - in progress, near finish
- AI module for Workflow Orchestrator - in progress

Network eAcademy AI track (not yet official)

AI for network automation

perfSONAR/PMP data analysis

Survey: Intelligent Networks - The Rise of Generative AI in Network Management



Study methodology

The use of GenAI in NM analysed across traditional ISO **FCAPS** categories (**F**ault, **C**onfiguration, **A**ccounting, **P**erformance, **S**ecurity) (deliberately **less focus on security**)

Analysed **peer-reviewed scientific papers and influential preprints** (2018 - June 2025) that explicitly employ **generative models** for network-management tasks.

Studies were retained if they:

(i) reported quantitative results on real or emulated networks,

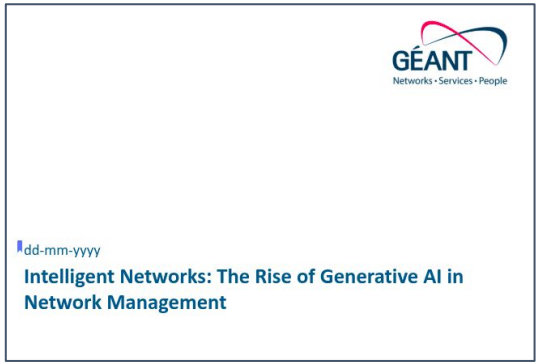
(ii) targeted at least one FCAPS function, and

(iii) offered code, data or methodological clarity sufficient for reproduction.

Total: 57 scientific references

The second part: the analysis of both **open source and commercial tools within the fields of AIOps** (Artificial Intelligence for IT Operations) and network performance management. **39 tool ref**.

Research team: prof Sonja Filiposka (UKiM), Dimitris Pantazatos (NTUA), Vincent Burkard (FAU), Claudia Torres Pérez (i2Cat/UPC)



The most widely explored families of generative models (1)

- **Large-Language Models** (**LLMs**) that transform text or code sequences;
  - **Prompt engineering** - involving carefully crafted input queries that guide the model's behaviour without modifying its internal configuration.
  - Using **fine-tuning**, one can retrain the model using domain-specific examples, including configurations, annotated incident logs, or policy documents.
  - A third strategy is **Retrieval-Augmented Generation (RAG)**, which integrates the LLM with an external knowledge source, such as a network documentation repository, logs, trouble tickets etc.
- **Generative Adversarial Networks** (**GANs**) - adversarial framework involving two neural networks: a generator, which produces synthetic data, and a discriminator, which attempts to distinguish generated data from real examples.
  - GANs have been widely adopted for **generating synthetic traffic data**, **simulating rare fault conditions**, and producing labelled intrusion examples for training security systems.

The most widely explored families of generative models (2)

- **Variational Autoencoders** (**VAEs**) that embed network states into smooth latent spaces for anomaly detection or configuration search; and
  - VAEs have found application in **modelling the distribution of valid configurations**, **detecting anomalies** via reconstruction error, and **generating intermediate network states** for smooth policy transitions
- **Diffusion Models** function by gradually adding noise to input data through multiple steps, then learning to reverse this process to generate structured outputs..
  - In network operations, diffusion models show promise in **synthesising configurations** that meet specific performance or QoS constraints and generating policies under operational constraints.
  - More computationally expensive than other models
- **Hybrid pipelines**

Research on the use of AI for Network Management

www.geant.org

Fault management

FCAPS

Fault management aims to detect, diagnose, and resolve problems with minimal impact on services.

Networks and systems are generating increasingly complex and multi-modal data, the ability to detect, diagnose and respond to faults quickly and accurately is essential to maintaining service reliability.

Opportunities for GenAI in fault management:

Semantic Log Interpretation

LLMs interpret and explain multi-modal logs to assist in identifying root causes

Predictive Fault Detection & Forecasting

Generative models anticipate faults and evaluate risk without affecting live systems

Synthetic Fault Pattern Generation

GANs generate rare or edge-case fault data for training robust detection models

Closed-loop Fault Management

AI-driven systems that detect, diagnose, and suggest remediation steps autonomously

Research papers - use cases

FCAPS

- Use transformer-based LLMs to **interpret unstructured log and telemetry data**. These models are used not only for fault detection but also for explainability
- Leverage GANs and VAEs to **generate synthetic fault scenarios**. These models help mitigate the problem of imbalanced fault datasets, improving the robustness of classifiers in low-frequency failure conditions.
- Integrate GenAI with **simulation environments**, building **digital twin systems**

- Reported **high anomaly detection accuracy**, **enhanced prediction** accuracy, **improved performance** - all that **in simulated environments**, **not integrated in the operational workflows**.

Strengths and Challenges

FCAPS

Strengths/promises:

- enhanced explainability, especially when using transformer-based LLMs that can interpret network logs
- proactive fault prediction
- research results are promising, e.g., 97-98% anomaly detection accuracy in some studies

Challenges

- Tests in real settings
- Operational reliability of synthetic data
- Integration complexity
- Latency and scalability.

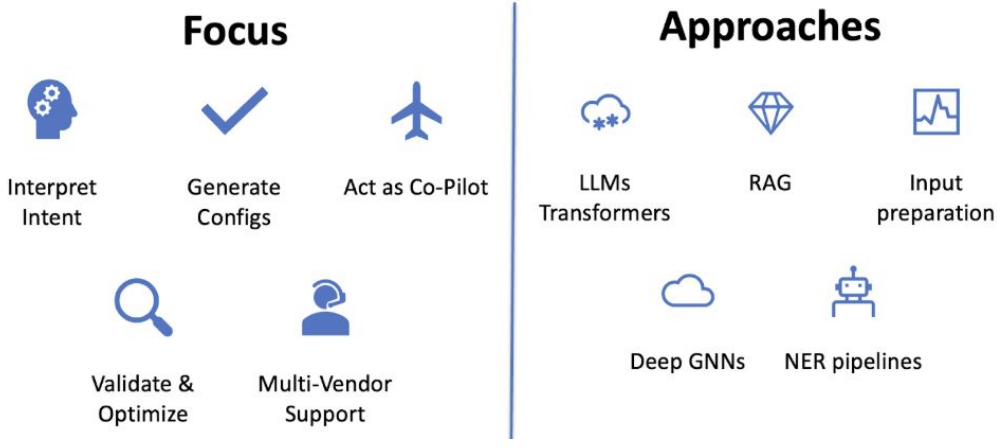
Configuration management

FCAPS

The largest number of research papers (apart from security)

Natural approach for GenAI (also seen in code development): By interpreting high-level input such as natural language or policy **intent**, generative models have the potential to support network engineers in producing accurate, consistent, and standards-aligned configurations.

Opportunities for GenAI in configuration management



Research papers - use cases

FCAPS

Some examples:

- *Generative AI for Low-Level NETCONF Configuration Management* combines LLMs with structured YANG models and natural language to **generate NETCONF-compliant XML configurations**
- The use of LLMs in combination with **standardised TM Forum and ETSI templates** to **transform high-level policy intent into complete configurations**
- **Co-pilot systems**, implement RAG to fetch relevant examples or prior configurations from a knowledge base and then use the LLM to generate context-aware suggestions
- *What Do LLMs Need to Synthesize Correct Router Configurations?* presents a Verified Prompt Programming approach, which combines **high-precision natural language prompts with post-generation validation**

Strengths and Challenges

FCAPS

Strengths:

- Simplifying and accelerating network configuration workflows
- Ability of models, especially fine-tuned LLMs and RAGs, to translate high-level, human-friendly inputs/**intents** into actionable configurations for real-world network devices
- Configuration **validation**
- Can be used in various **multi vendor settings** (IP networks, 5/6G,...)

Challenges:

- LLMs can produce syntactically correct but semantically flawed configurations
- Accuracy and reliability of LLMs still vary based on the complexity of the input
- Models struggle with edge cases, or uncommon syntax variations
- Generalisation and reliability
- Compliance and auditability
- Explainability and trust

A consistent and important design principle: the emphasis on **keeping humans in the loop** as a deliberate architectural choice. Incorporate **co-pilot models** that support, rather than replace, engineers.

Accounting Management

FCAPS

Accounting refers to the monitoring, measuring, and analysing resource usage across a network, typically for billing, cost optimisation, auditing, and service-level enforcement.

No peer-reviewed studies that explicitly focus on using GenAI for accounting tasks in a networking context have been identified.

However, there is some related work in:

- financial accounting,
- automated billing systems, and
- predictive analytics for revenue forecasting

Which demonstrates the potential of LLMs, VAEs, and GANs for generating synthetic transaction data, detecting anomalies, and producing explainable summaries.

Performance Management

FCAPS

Performance management in networking systems focuses on ensuring the availability, responsiveness and efficiency of network resources.

Opportunities for GenAI in network performance management

Traffic Forecasting

Predict traffic patterns using time-series generative models

Example: Avoid congestion before it occurs

QoS Metric Inference

Infer latency, jitter, throughput directly from traffic traces

Example: Auto-monitor SLA performance

Adaptive Resource Allocation

Use GANs in RL frameworks to recommend bandwidth/resource changes

Example: Real-time load balancing

Model Lifecycle Support

Trigger retraining based on GenAI-simulated drift

Example: Keep ML models tuned without constant re-labelling

Research papers - use cases

FCAPS

Some examples:

- Diffusion models can **simulate mobile user behaviour** and **design contracts** that optimise wireless service delivery under QoS constraints
- *Deep-Q* - deep generative model to directly **infer QoS metrics** from raw traffic traces
- *GAN-powered Deep Distributional Reinforcement Learning* - **dynamic resource allocation in network slicing**. GAN-generated synthetic samples can accelerate policy learning, leading to faster convergence and more stable performance even under traffic fluctuation.
- *OpticGAI* targets optical networks, combining generative learning with deep reinforcement techniques to **optimise complex problems like routing and spectrum allocation**.

Strengths and Challenges

FCAPS

Strengths:

- Traffic prediction, network adaptation, and optimisation under dynamic network conditions
- Anticipate traffic loads, simulate user behaviour, infer quality metrics, and support intelligent decision-making frameworks.
- Proactive performance tuning
- Support complex decision-making under uncertainty

Challenges:

- high-quality training data, both in terms of representativeness and structure.
- computational cost

Security Management

FCAPS

Large body of AI/ML use in cybersecurity for anomaly/attack detection

Opportunities for GenAI use in Security management:

- synthetic attack generation
- prompt-based rule generation (model can write rules that define protocols, offsets, byte patterns, flow direction, thresholds and metadata in Snort/Suricata engine's native, keyword-rich syntax)
- data augmentation (using GANs)
- unsupervised anomaly detection (using VAE)
- Similar uses like fault management

AI Tools for Network Management

www.geant.org

Vendor solutions

"All in one" approach - covering (almost) all FCAPS areas

Tailored for the products of the company

Common features:

- Natural language interaction,
- Configuration assistant
- Script generation
- Troubleshooting assistant
- Automated incident summarization
- Anomaly detection

Cisco AI Network Analytics

MIST AI

watsonx

FortiAI

ServiceNow ITOM

XENONSTACK
A Stack Innovator

ATERA

NetBrain Chatbot

DX Operational Observability

Open-source solutions and frameworks

Not tailored for NM, but can be used/adapted

Covering less FCAPS areas

Also analysed additional 15 security related AI tools

Common features:

- Script generation,
- natural language to device configuration
- LLM interface, Model Context Protocol (MCP)
- autonomous problem-solving

LangChain

Pydantic AI

AutoGen

CAI CYBERSECURITY AI Bug bounty-ready AI

Adversarial Robustness Toolbox

Agent Development Kit

Conclusions

Configuration management and security management are the most advanced in adoption and capability. Fault management is following closely behind.

Commercial solutions generally provide dedicated vendor support and are closer to the agentic AI, whereas open-source tools depend more on community support and require more customization.

Many promises, but some challenges remain:

- Validity and safety
- Explainability
- Operational constraints such as latency, integration with other systems, and compliance
- How justified is to rely on synthetic data?

This study is the basis for further WP6 actions in the Network eAcademy and incubators

Appendix

Some more recent results

Two studies of the use and the adoption of the GenAI in IT and business in general:

June 25: Yegor Denisov-Blanch, Stanford, Does AI Actually Boost Developer Productivity? (100k Devs Study)

The analysis of 100.000 developers, millions of commits, billions lines of code, 80% private repos
https://www.youtube.com/watch?v=tbDDYKRFjhk&list=WL&index=29

July 25: State Of AI In Business 2025, MIT Project NANDA

Methodology: interviews, surveys, and analysis of 300 public implementations
https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf

Stanford study - AI generated code needs re-work

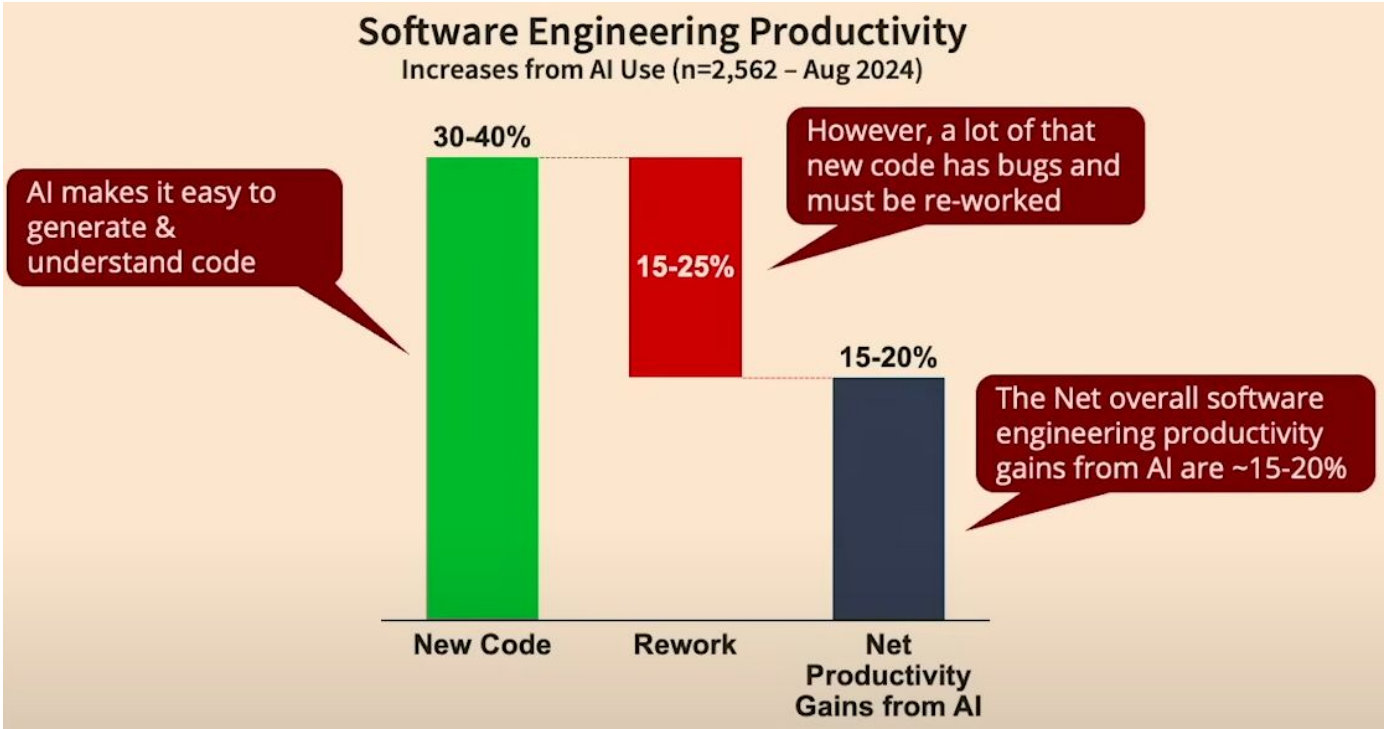


Image taken from: Does AI Actually Boost Developer Productivity? (100k Devs Study) - Yegor Denisov-Blanch, Stanford: https://www.youtube.com/watch?v=tbDDYKRFjhk&list=WL&index=29

Stanford study - Gains are larger with simple greenfield tasks

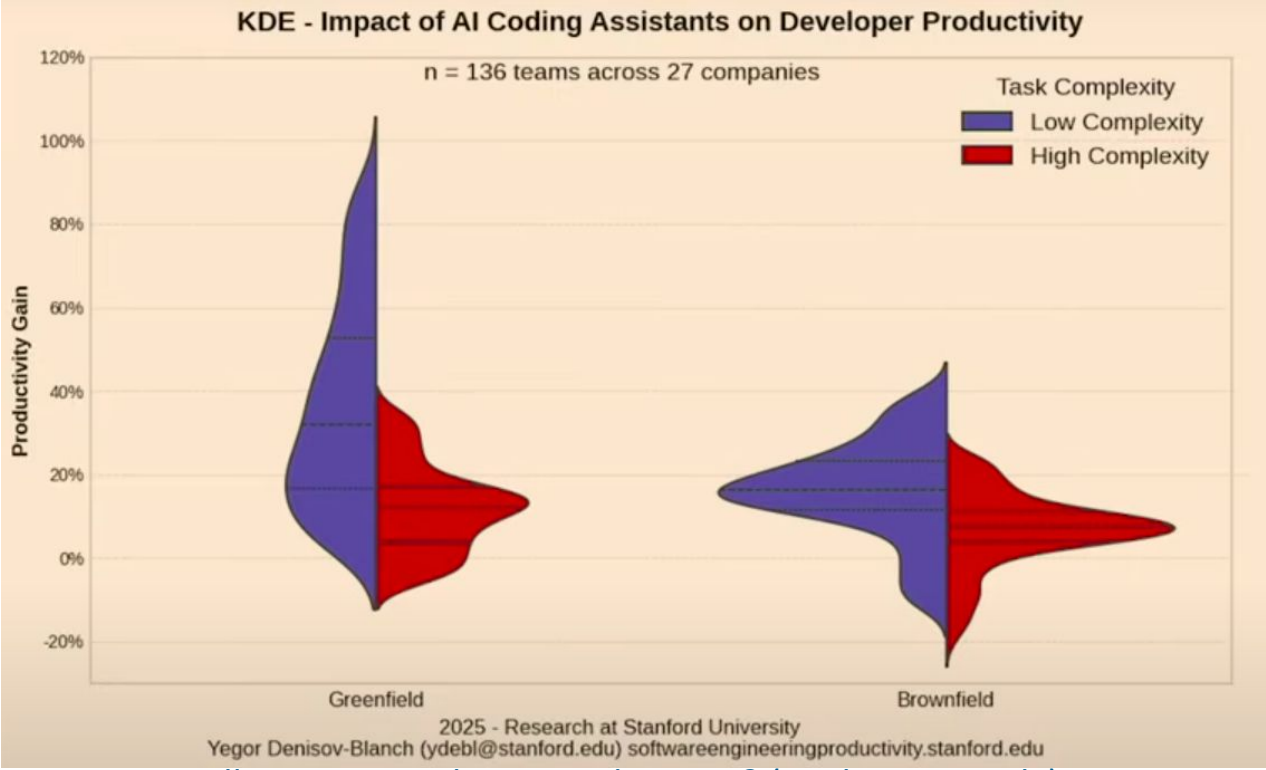


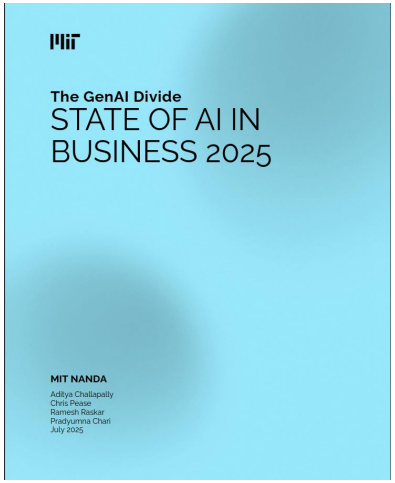
Image taken from: Does AI Actually Boost Developer Productivity? (100k Devs Study) - Yegor Denisov-Blanch, Stanford: https://www.youtube.com/watch?v=tbDDYKRFjhk&list=WL&index=29

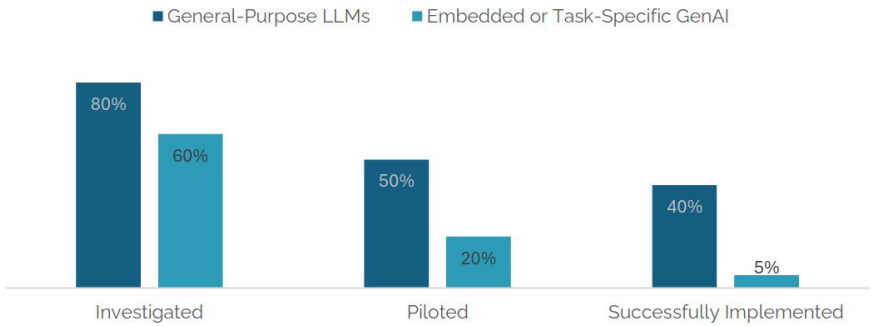
MIT NANDA study (1)

GenAI divide - 95% of organizations are getting zero return from GenAI

Key takeaways:

- Adoption is high, but disruption is low. Seven of nine sectors show little structural change (But Telco is in the other two!). Enterprises are piloting GenAI tools, but very few reach deployment
- Only 5% of custom enterprise AI tools reach production. Chatbots succeed because they're easy to try and flexible, but fail in critical workflows due to lack of memory and customization.





MIT NANDA study (2)

Key takeaways (2):

- While official enterprise initiatives remain stuck on the wrong side of the GenAI Divide, employees are already crossing it through personal AI tools. Employee usage far outpaces official adoption.
- Key barrier: users resist tools that don't adapt, model quality fails without context
- Five Myths About GenAI in the Enterprise

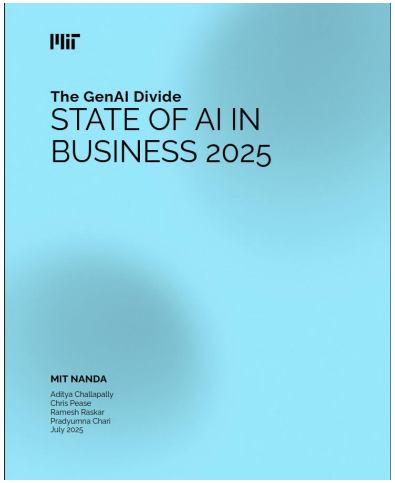
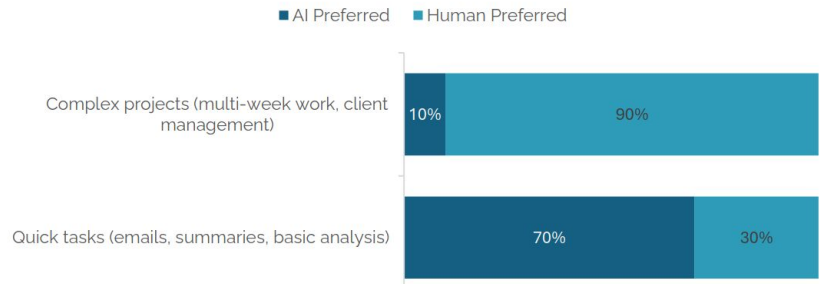
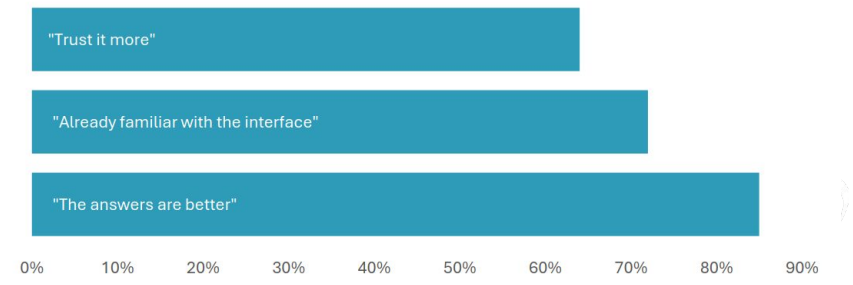


Exhibit: Perceived Fitness for High-Stakes Work
"Would you assign this task to AI or a junior colleague?"



User Preference Drivers: Generic LLM Interface vs. Integrated Tool



Thank You

Any questions?

www.geant.org