

# Comparing Notes on Trustworthy AI 2025

## Part 3: AI Agents in the Workplace: Human Labour, Collaboration, and Societal Impact

### *German version below*

**Host:** TUM Think Tank

**Date:** 24 July 2025

**Location:** TUM Think Tank, Munich

The third event shifted focus from the technical to the human: how do AI agents change the structures of work, not just individual tasks? What does responsible deployment look like when the impacts are social and organisational as much as technical?

### The panel

**Leif-Nissen Lundbak** works at the intersection of regulation, risk, and trust. He described the shift from the Oracle model of AI to the Genie model (delegation of tasks with autonomous return) and argued that the EU AI Act is too model-centric.

**Nicklas Lundbald** focuses on technology and governance. He reinforced the shift from fact-based to act-based trust, arguing that liability should move toward ecosystem-level accountability.

**Tobias Kampf** researches work structures and labour. He stated that AI agents fundamentally change workplace structures and still rely on human labour.

**Edyta Bogucka** studies multi-layered risk and misuse. She identified three layers of impact: model evaluation (hallucinations), human interaction (70% of failures are due to humans), and systemic impact.

**Nadine Stockinger** works on agents in industrial collaboration, highlighting their use in vehicle development to accelerate collaboration and reduce workload.

**Tina Weisser** focuses on process redesign, advocating for rethinking entire processes to increase transparency and shift leadership dynamics.

The panel was moderated by **Dalia Yousif**, doctoral candidate and research assistant at the TUM School of Social Sciences and Technology.

### Discussion

#### **How AI agents change work structures**

AI agents change the structures of work, not just tasks. Legal agents will automate repetitive tasks, impacting junior professionals most. Notably, 70% of failures stem from human factors, such as users giving AI meaningless tasks or using agents to fake productivity.

## **Human-AI collaboration models**

The panel emphasized wholesale process redesign over simple digitization. AI offers opportunities to shift power dynamics and increase transparency. Multi-agent environments will require new coordination roles for managing interactions between agents and humans.

## **Governance, responsibility, and organisational practice**

Governance models must clearly allocate responsibility, especially around data inputs. Sector-specific governance is preferred over one-size-fits-all solutions. Professional judgement must be maintained by overseeing agent outputs.

## **Risks, ethics, and employee well-being**

Probabilistic automation introduces unpredictability. The focus should be on protecting workers' rights rather than preserving current jobs. Ethical risks increase with multi-agent interactions, requiring frameworks like constitutional AI.

## **Key takeaways**

The most successful approach to AI agents in the workplace requires:

- Holistic redesign rather than simple digitisation of existing processes.
- Focus on worker protection rather than job preservation.
- Sector-specific governance rather than one-size-fits-all solutions.
- Active management of social dynamics alongside technical implementation.
- Experimentation culture while maintaining ethical responsibility.
- Gradual trust building through transparent communication.

# Comparing Notes on Trustworthy AI 2025

## Teil 3: KI-Agenten am Arbeitsplatz: Menschliche Arbeit, Kollaboration und gesellschaftliche Auswirkungen

**English version above**

**Gastgeber:** TUM Think Tank

**Datum:** 24. Juli 2025

**Ort:** TUM Think Tank, München

Das dritte Event verlagerte den Fokus vom Technischen auf das Menschliche: Wie verändern KI-Agenten die Strukturen der Arbeit, nicht nur einzelne Aufgaben? Was bedeutet verantwortungsvoller Einsatz, wenn die Auswirkungen so sehr sozial und organisatorisch wie technisch sind?

### Das Panel

**Leif-Nissen Lundbak** beschrieb den Übergang vom Oracle-Modell der KI zum Genie-Modell und argumentierte, dass der EU AI Act zu modellzentriert ist.

**Nicklas Lundbald** betonte die Verschiebung von faktenbasiertem zu handlungsbasiertem Vertrauen und forderte Ökosystem-Ebenen-Verantwortlichkeit.

**Tobias Kämpf** erklärte, dass KI-Agenten Arbeitsplatzstrukturen fundamental verändern und weiterhin auf menschliche Arbeit angewiesen sind.

**Edyta Bogucka** identifizierte drei Auswirkungsebenen: Modellbewertung, menschliche Interaktion (70% der Fehler durch Menschen) und systemische Auswirkungen.

**Nadine Stockinger** berichtete über den Einsatz von KI-Agenten in der Fahrzeugentwicklung zur Beschleunigung der Zusammenarbeit.

**Tina Weisser** forderte, Prozesse neu zu denken, statt sie nur zu digitalisieren, um Führungsdynamiken zu verändern.

Das Panel wurde moderiert von **Dalia Yousif**, Doktorandin an der TUM School of Social Sciences and Technology.

### Diskussion

#### Wie KI-Agenten Arbeitsstrukturen verändern

KI-Agenten verändern Arbeitsstrukturen, nicht nur Aufgaben. Im Rechtsbereich werden Aufgaben wie NDA-Prüfung automatisiert; Junior-Rechtsfachleute sind am stärksten betroffen. Branchenübergreifend variiert die Adoption stark nach nationaler Arbeitskultur. 70% der Fehler entstehen durch menschliche Faktoren - Nutzer, die KI sinnlose Aufgaben geben, oder Agenten, die zur Vortäuschung von Produktivität eingesetzt werden.

## **Mensch-KI-Kollaborationsmodelle**

Das Panel lehnte einfache Digitalisierung ab und forderte ganzheitliche Prozess-Neugestaltung. KI bietet Möglichkeiten, Macht zu verschieben und Transparenz zu erhöhen. Multi-Agenten-Umgebungen brauchen neue Koordinationsrollen. Adoption ist ebenso ein soziales wie ein technisches Problem.

## **Governance, Verantwortung und organisatorische Praxis**

Governance-Modelle müssen Verantwortung klar zuweisen. Sektorspezifische Governance wird gegenüber Einheitslösungen bevorzugt. Bei BMW wird jede KI-Bereitstellung von einem Rat genehmigt. Professionelles Urteilsvermögen in rechtlichen Entscheidungen muss durch Überwachung der Agenten-Ausgaben bewahrt werden.

## **Risiken, Ethik und Mitarbeiter Wohlbefinden**

Probabilistische Automatisierung führt Unvorhersehbarkeit ein. Der Fokus sollte auf dem Schutz der Arbeitnehmerrechte liegen, nicht auf dem Erhalt aktueller Arbeitsplätze. Ethisches Risiko steigt, wenn mehrere Agenten interagieren. KI kann bestehende Ungleichheiten verschärfen, wenn nicht mit Inklusivität im Sinn gestaltet.

## **Wichtigste Erkenntnisse**

- KI-Agenten sind kein vorübergehender Trend - sie verändern die Arbeitsstrukturen fundamental.
- 70% der Fehler entstehen durch menschliche Faktoren, nicht durch fehlerhafte KI.
- Multi-Agenten-Systeme erfordern neue Governance-Ansätze.
- Erfolg hängt von qualifizierten Mitarbeitern und deren Akzeptanz ab.
- Europäische Werte und Datensouveränität müssen in die Entwicklung eingebettet werden.
- Sektorspezifische Governance statt Einheitslösungen.
- Schrittweiser Vertrauensaufbau durch transparente Kommunikation.