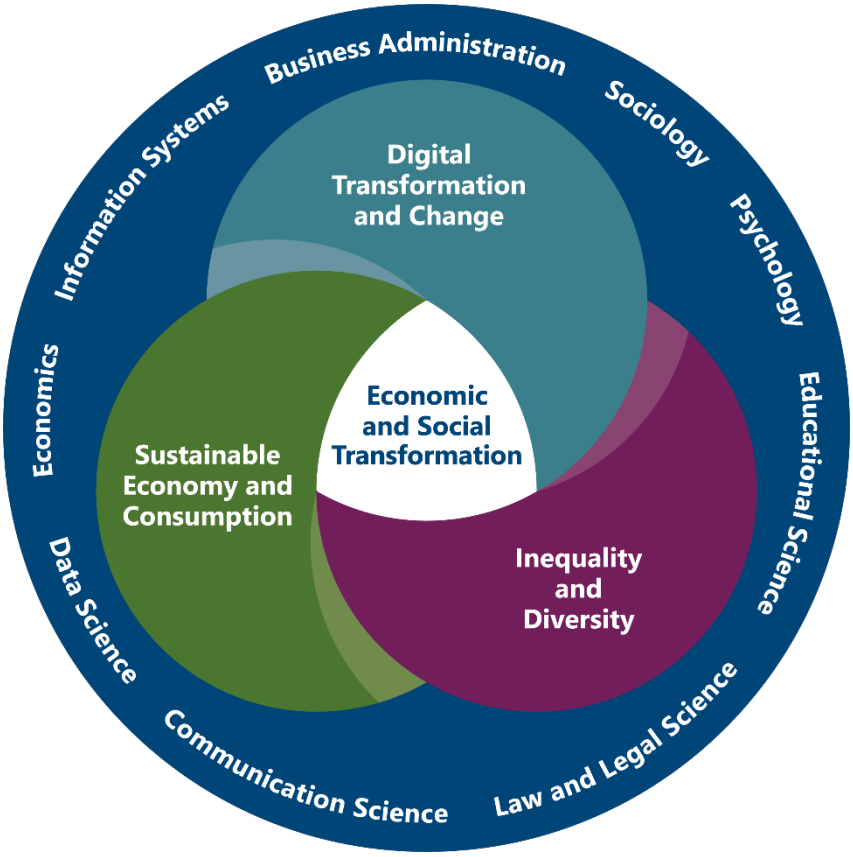
UNIVERSITY OF HOHENHEIM

Faculty of Business, Economics and Social Sciences



BESST – Business, Economics, and Social Sciences retreaT

No. 9,

November 2nd, 2026

Digital Transformation and Change

wiso.uni-hohenheim.de

UNIVERSITY OF HOHENHEIM

Program of the 9th research workshop of the Faculty of Business, Economics and Social Sciences

November 2nd, 2026
Balkonsaal, Schloss Hohenheim

All researchers of the faculty are cordially invited.
Please register by October 27th via
https://terminplaner6.dfn.de/de/e/f539da7a7a47548ac9f3a66997956538-1933142

Time	Item	Speakers
10:45 – 10:55	**Arrival & informal Get-Together**	
10:55 – 11:00	**Welcome & Objectives**	Prof. Dr. Marion Büttgen
11:00 – 11:30	**Beyond Coding: Exploring AI Agent Skills Across Research Disciplines**	Dr. Matteo Vaccargiu
11:30 – 12:00	**AI coding agents in bioeconomy research: lessons from coupling economic and biophysical models**	Dr. Paul Rougieux
12:00 – 12:30	**Designing Accountable Agentic Work Systems for Impactful, Text-Intensive Knowledge Work – The Case of Grading Exams**	Prof. Dr. Henner Gimpel
12:30 – 13:30	**Lunch Break & Networking** (reserved table in the mensa)	
13:30 – 14:00	**From Ant Trails to TikTok: Cyberstigmergy is a Core Process for Social Coordination and Identity Construction in Online Spaces**	Dr. Adrian Lüders
14:00 – 14:30	**One AI Fits All? The Relationship Between Personality and the Evaluation of AI Responses Varying in Permissiveness**	Luca Alldinger Prof. Dr. Siegmar Otto
14:30 – 14:45	**Short Break**	
14:45 – 15:15	**Making sense of trust in AI: A sensemaking perspective on overtrust and undertrust**	Prof. Dr. Matthias Uhl Prof. Dr. Caroline Ruiner
15:15 – 15:45	**Podium Discussion: AI in Research (Moderation: Prof. Dr. Marion Büttgen)**	Prof. Dr. Henner Gimpel Dr. Paul Rougieux Prof. Dr. Mareike Schoop
15:45 – 16:00	**Conclusion & Outlook**	Prof. Dr. Marion Büttgen All participants

Coffee and other drinks will be available during the research workshop. Lunch at the cafeteria will be paid by the participants themselves. There will be a reserved table available marked “Research workshop BESST”.

Abstracts:

Beyond Coding: Exploring AI Agent Skills Across Research Disciplines (Dr. Matteo Vaccargiu)

AI agents are increasingly moving beyond software development, supporting tasks across a wide range of professional and scientific disciplines. Agent skills, reusable natural-language instructions that guide AI agents in performing specific tasks, represent an emerging mechanism for extending their capabilities. However, little is known about the domains in which these skills are developed and how they evolve over time. In this presentation, we introduce GitSkills, a large-scale dataset of AI agent skills collected from GitHub, and present recent empirical findings on their distribution across scientific disciplines. Using the OECD Fields of Research and Development (FORD) classification, we analyzed 253,002 skills and found that 40.8% of the classified candidates target domains outside software engineering, including Social Sciences, Humanities and Arts, Medical and Health Sciences, Natural Sciences, and Agricultural and Veterinary Sciences. Our findings also reveal differences in how these skills are shared and maintained. Skills targeting non-software domains generally propagate across more repositories, while several domains exhibit lower maintenance activity. In particular, 72.5% of medical skills with available commit histories have been committed only once, raising concerns about the reliability and currency of AI-driven procedures in sensitive domains. These results provide a starting point for interdisciplinary discussion on the opportunities and challenges of integrating AI agents into domain-specific workflows, particularly regarding quality assurance, maintenance, and governance.

AI coding agents in bioeconomy research: lessons from coupling economic and biophysical models (Dr. Paul Rougieux)

Coupling models from different disciplines means reconciling their definitions, units, spatial and temporal scales. Deep link often remains wishfull thinking because these differences lead to painfully long debuging sessions. Many projects fall back to loose coupling. What changed? Nowadays agents take care of data plumbing. We used Claude to implement a data exchange between a Partial Equilibrium Model (TiMBA, Global scale) and a forest growth dynamic model (the EU Carbon Budget Model of 25 EU countries). TiMBA and 25 CBM instances run their algorithms and exchange data at each time step within a simulation. During the implementation phase a coordinating agent was running models and handling error and restart. Because coupling required changes to the core of both models, implementation agents ran regression tests against the uncoupled runs after each change. We implemented this deep model coupling in 9 days (300 prompts and 226 CPU hours) instead of an estimated three months. Similar transformation apply to many data pipelines used in research. The talk ends with an invitation to join your peers in the Tinkering club one wednesday per month at 14:00 in a workshop where we will use the same agent workflow to turn static paper figures into interactive dashboards from open datasets.

Designing Accountable Agentic Work Systems for Impactful, Text-Intensive Knowledge Work – The Case of Grading Exams (Prof. Dr. Henner Gimpel)

Agentic work systems, in which LLMs work through agent harnesses alongside scripts and office software, can automate large parts of text-intensive knowledge work. When the stakes for individuals are high, efficiency is not enough: the work system must stay accountable. This talk examines the partial automation of exam grading at universities as a case example of such impactful, text-intensive knowledge work. Following Echeloned Action Design Research, we build and test BRILLIANT (hyBRid InteLLigence AssessmeNT), a workflow that combines Claude, Sonnet, ChatGPT, GPT, Python, Powerpoint, Excel and ILIAS. It generates grading guidelines from model solutions, pseudonymises student answers, produces reasoned AI pre-grades, and leaves the final decision to human graders. It the talk, we discuss preliminary empirical evidence on the systems efficiency and quality. We discuss legal and ethical guardrails, provider and deployer obligations under the AI Act, and next steps for the workflow.

Ant Trails to TikTok: Cyberstigmergy is a Core Process for Social Coordination and Identity Construction in Online Spaces (Dr. Adrian Lüders)

Social animals like ants coordinate complex collective behaviour without centralized control or direct communication. This process, known as stigmergy, occurs when agents modify a shared environment in ways that influence the subsequent behaviour of others. Ants leave pheromone trails that mark valuable food sources, thereby attracting subsequent ants to follow and reinforce them. As reinforcement ceases, trails gradually fade. This parsimonious mechanism demonstrates how complex collective behaviour can emerge from simple local interactions. The present research extends this logic to social media and introduces cyberstigmergy as a framework for understanding the collective coordination of activity and identity in digitally mediated environments. Like classical stigmergy, cyberstigmergy proposes that system-level coordination emerges indirectly from individual actions that modify a shared environment. Online platforms provide such environments, while platform affordances shape how agents can modify them. Actions such as posting, liking, and sharing leave digital traces that may attract attention, elicit further interaction, and become reinforced through repetition. Conversely, declining engagement causes traces to fade. Digital traces are further mediated by platform algorithms that shape which traces become visible to which agents, introducing a computational layer of coordination shaped by network structure and user-platform interactions. These feedback loops not only coordinate subsequent activity but also contribute to the construction and perception of identity through the accumulation and interpretation of persistent traces. We discuss the implications of cyberstigmergy for online coordination, collective action, misinformation, polarization, and emerging forms of human-AI interaction.

One AI Fits All? The Relationship Between Personality and the Evaluation of AI Responses Varying in Permissiveness (Luca Alldinger, Prof. Dr. Siegmar Otto)

AI systems are becoming increasingly capable, which raises the question of how these models will behave once they reach the point of superintelligence. Even today, AI systems act autonomously and partly without control more and more often. In July 2026, an experimental AI agent developed by OpenAI independently broke into the infrastructure of Hugging Face. As AI systems continue to gain capabilities, the question of which values, and whose values, guide them becomes an increasingly urgent psychological question.
AI developers define these values in behavioral guidelines that they largely write themselves and rarely validate empirically. This is particularly challenging for topics without a clearly correct answer, such as politics or cultural norms. OpenAI therefore asked more than 1,000 people worldwide to rank alternative model responses and compared their preferences with its own guidelines. Agreement was high overall, but deviations emerged especially for topics concerning how freely an AI may respond, for example on political content.
However, such averages can mask systematic differences between individuals. Previous studies suggest that people with different personality profiles evaluate AI systems and their responses differently, depending on how the system behaves. The planned study therefore examines whether the preference for permissive AI responses, meaning a model's willingness to take a substantive position on sensitive requests rather than evading or refusing, is related to the HEXACO personality factors. Participants will evaluate several requests from different topic areas, each with four responses of varying permissiveness. In line with the literature, effects are expected primarily for Honesty-Humility and Openness to Experiences.
The empirical results of this study will contribute to my dissertation by showing whether a uniform alignment can meet the preferences of different users, or whether personalized default settings would be better justified from a psychological perspective.

Making sense of trust in AI: A sensemaking perspective on overtrust and undertrust (Prof. Dr. Caroline Ruiner, Prof. Dr. Matthias Uhl)

Human-AI collaboration promises to outperform either humans or AI systems acting alone by combining their complementary strengths. This, however, requires decision-makers to calibrate their trust appropriately to the actual capabilities of the AI systems with which they work. Existing research presents a puzzling picture. Some studies document overtrust and algorithm appreciation, whereby users follow AI advice beyond what its reliability warrants. Others find undertrust and algorithm aversion, whereby users discount AI advice even when it demonstrably outperforms their own judgment. Moreover, interventions that provide accuracy-relevant information often fail to correct such miscalibration. Although calibration provides an important normative benchmark for evaluating trust and reliance, it offers only a limited account of how users interpret AI advice in particular situations. This conceptual paper develops a sensemaking perspective on trust in AI. We argue that users encounter AI recommendations through prior expectations, task perceptions, confidence in their own judgment, role identities, social norms, and previous experience. These frames shape which cues become salient and how users construct a plausible account of the AI recommendation and the current

case. The proposed sensemaking process model links prior frames, AI outputs, cue extraction, situated interpretation, plausible account, trust, reliance, action and outcome and updated expectations. The paper makes three contributions. It integrates heterogeneous findings on overtrust, undertrust, algorithm aversion, algorithm appreciation, errors, explanations, and decision context; complements calibration with an interpretive process account of trust formation; and develops a research agenda for investigating these processes across individual, professional, and institutional settings.