Creative Intent and Reflective Practices for Reliable and Performative Human-AI Systems

Uta Wilkens, Ambrose Field

1. Introduction

"AI will probaly most likely lead to the end of the world, but in the meantime, there'll be great companies" (Sam Altman, OpenAI Co-Founder & CEO). Even though this statement supposed to be polemic and was later characterized as "partly in jest" (Shin 2023), it is the founder behind ChatGPT who signed the open letter on AI risks for humanity: "Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war (https://www.safe.ai/statement-on-ai-risk#open-letter) – a declaration among leading CEOs of the US digital industry and researcher in computer science that was published in June 2023.

"The only things in my life that compatibly exists with this grand universe are the creative works of the human spirit" (Ansel Adams, Feb. 20, 1902 – Apr 22, 1984). The American environmentalist and landscape photographer Ansel Adams died exacly one year before Altman was born. Both are convinced that Big Sur, an area in California with lush vegetation, would be the best place to live on earth. Both reflect on the universe from their very different professions and let us think about human creative intent as the most valuable source of sustaining humanity in future (Dewey 1934). Nowadays it is supposed to be a mastery of sophisticated technology. As Maeda (2002, p. 39) argues we risk that technology becomes "an end in itself in society and industry" when design is not driven by human creative intent. Since ancient Greek creative intent is described as the way how human beings use and combine their contextual knowledge and experiences to find new solutions (Dewey 1934). Novelty of the outcome was not the most important characteristic in the origin writings, rather the underlying process of generating solutions while using all senses.

The world could recently observe that it was indigenous knowledge of how to survive in the jungle when four children between 13 years and 11 months managed to stay alive for 40 days in the Columbian rainforest after a plane crashed. If it is human creative intent augmented but not replaced or disturbed by technology there might be a way to cope with the risk of technology and to use it for better

DOI: https://doi.org/10.30844/wgab_2023_5

life and better working conditions (Fischer 2018) even in less developd and priviliged countries. The prerequisite is to put human creative intent and reflective practices in the center of technology design.

The movement of Industry 4.0 neglected the human-side of generating solutions while giving high attention to autonomous systems communicating with each other and regulating challenges by smart sensor technology (Lasi et al. 2014). The vision for manufacturing was that "products control their own manufacturing process" (Lasi et al. 2014, p. 239). With respect to digital services the idea was that technology takes the responsibility for regulating critical situations, e.g. for higher safety in autonomous driving (Fagnant/Kockelman 2015). In the first wave of Industry 4.0 there was a belief that sensor technology could work autonomously and provide the necessary sensory input of the system. Expectations related to this vision seemed to be too high; a revolution of autonomous systems did not take place. From a technical point of view the reason was attributed to underdeveloped connectivity and interoperability due to missing standards across firms and industries (Jepsen et al. 2020). But there is also the critical reflection that sensor technology is too limited to manage critical interfaces and to provide reliable and resilient solutions in unforseen situations. The human sense and contextualized knowledge for keeping systems work was missing in the technology design.

Meanwhile, there is a vision of Industry 5.0 aiming at higher human-centricity in terms of collaborative hybrid systems. This movement aims to keep the human in the loop by design and to elaborate on semi-autonomous instead of autonomous systems (Nahavandi 2019). Even though the expression semi-autonomy is considered as characteristics of robots it includes both sides, the autonomy of the human actor and the autonomy of the non-human actor which are interrelated on system level.

We want to elaborate on this human side and ask how it can serve as most valuable source of reliable systems. Our core argument is that human-centricity is a necessary condition to reach and sustain the intended outcomes of AI-based systems in terms of safety, quality, accuracy and reliability. The reflection on human creative intent and how to build semi-autonomous systems around is a matter of avoiding the risks for humanity expressed in the recent declaration of leading AI experts. The aim of this contribution is to explain why human creative intent as an input and throughput factor and reflective practices for a continuous interaction between AI and humans is a crucial point to sustain semi-autonomous systems and gain solutions on a relational basis – a circumstance demanding new methods in co-design and co-creation.

In the next sections we outline the relational constructs keeping systems work from a theoretical point of view. We will illustrate the meaning and relevance with the help of five use cases of high relevance for a bright future of our societies. The following discussion gives emphasis to design principles of how to keep human creative intent and performative interaction in the center of technology development. Finally we give an outlook on necessary research methodology for the future research agenda.

2. Relational constructs of human-AI systems

2.1. Creative intent as individual relational practice for system development

Nowadays intent is a construct of contemporary political psychology with emphasis on interest-driven outcomes that can be explained by contextual factors of the interest groups. Among these scholars there is stil a reference to the roots in ancient Greek phylosophy (Neblo 2007). We trace back to these roots as it allows us to integrate a broader discourse of relevance for systems in work and business instead of primarily considering a political agenda.

“Intent” is a concept coined by the ancient Greeks, later developed through the work of John Dewey (1934) and applied across industrial contexts by Rogala et al. (2020). The relevance of human creative underpinnings for technology development and especially the risks of their absense was outlined by Maeda (2002). The construct brings together human contextual and environmental factors with the processes of creating and making, and as such, has high relevance in today's complex systems design. Intent requires realisation through performative actions. It is a flexible, dynamic construct that can be described as key to driving developmental progress towards solutions. However, it is often far from a perfectly formed solution, it is neither irrelevant to finding one.

Intent, as a concept, is a key part of creativity (Maeda 2002): It is a quality that becomes modified in response to evolving solutions, it reflects specific, contextualised knowledge and it shapes the onwards flow of decision making and enquiry. In creative domains, intent forms a guiding marker for creativity. It is a synthesis of human analytical, emotional and relational processes. Creative intent is a driver for curiosity – the 'what if questions' and at the same time way to it’s answer. It describes creativity as a process and not as a novel outcome.

Referring to the system level and conceptualizing intent as a relational construct the contextualization plays a pivotal role. Rheinberger (1997) argues in correspondence with environmental biology, that the limitations of experimental process are a function of the application context. In his concept of “experimental system” Rheinberger (1997) describes a space where contextual considerations create networked thought, rather than a more abstracted and procedural test/experiment/evaluate workflow common to lab procedures. It is intent that helps steer a path towards outcomes – forming a guiding framework for contextualised experiment within a complex system.

DOI: https://doi.org/10.30844/wgab_2023_5

Among the German speaking research community there is a synonomous construct which is especially considered as relevant in demanding work settings: “Kompetenz” is defined as the ability to act and interact against the background of new tasks and demands and to find coping patterns and novel solutions within continuously changing process. The meaning of “Kompetenz” is somehow routed in the German vocational system (Erpenbeck 2002) and addresses the meta-ability for self-organization in unfamiliar situations that require to activate and combine cognitive, methodological and social abilities to solve problems as an issue of performative action (Erpenbeck 2002; Heyse/Erpenbeck 2007; Erpenbeck et al. 2017; Wilkens et al. 2006). E.g. the multi-level model proposed and validated by Wilkens/Sprafke (2019) builds on coping with complexity, self-reflection on action, combination and cooperation within and across organizations in order to develop situated solutions within a contextualized learning process on system level. "Kompetenz" requires a deep sense for classifying situations and needs, understanding technology, material and service not just on a cognitive level testified with certificates of qualification but also with all human senses and in a socialized manner (Erpenbeck 2002). This allows to not just operate according to task descriptions but to reflect on meaningful solutions (see figure 1, left hand side).

Elaborating on these synonomous constructs it becomes obvious that their relevance rises when there are systems with high dynamics without a pre-defined best way or solutions but rather depending on relibale performative interactions. This is the case for AI-based systems, especially when the technology gains agentic character and operates in an autonomous manner (see Kaartemo/Helkkula 2018 and next section).

2.2. Understanding the relational character of AI

"Artificial Intelligence (AI)" is nothing more than a terminology which has no parallels to individual intelligence. In its origin the ability of a software could also have been named as computational detection of patterns with the help of neural networks (Barthakur 2023; Wilkens 2020). There is no general definition for AI and the meaning evolves with new generations of technology (Launchbury 2017; Xu 2019). The current state of the art in technology development is artificial general intelligence in the meaning of “intelligent agents that will match human capabilities for understanding and learning any intellectual task that a human being can” (Fischer 2023, p. 1). The generative AI ChatGPT gives an example that goes into this direction (Barthakur 2023, see also section 2.4). Considering cases with AI use in the workplace it is more algorithms dedicated to solve specific and pre-defined problems where the technology was pre-trained for and fine tuned on the basis of a mass of data. This gives “machines the ability to reason and perform cognitive functions such as problem solving, object and word recognition, and decision-making” (Hashimoto et al. 2018, p. 70) in a specified field. These AI-based operations result from machine learning (ML) for detecting patterns in terms of supervised learning (detecting pre-labeled patterns, e.g. distinguishing between fault and correct products in production line), unsupervised learning (detecting relationships between data points, e.g. consumer preferences for a set of products) or reinforcement learning (algorithms aiming at a specific goal based on punishment and reward functions, e.g. gender equality in personnel selection) (Russell/Norving 2021). This goes beyond former processes of automation resulting from pre-programmed static commands as AI can learn from experiences and new data gathered from operational processes and thus act and develop autonomously (Haenlein/Kaplan 2019) without further human intervention (van Rijmenam/Logue 2020) as outlined in the origin Industry 4.0 senarios considering AI-based systems communicating with each other autonomously and regulating challenges by smart sensor technology (Lasi et al. 2014).

Considering this "autonomy" of the system from the perspective of learning and development it is important to note that the autonomy is not multi-directed. The underlying learning process is completely different from individual learning processes with their potential to learn and develop beyond and out-of-the-box while combining insights and experiences from different contexts (Wilkens 2020). Since an ML approach supports learning in only one direction, primarely on a single-loop basis of further optimizing the system but at the same times undermines double-loop learning options while doing so, Wilkens (2020) characterizes AI-based learning as a "double-edged sword". Autonomous AI systems cannot cope with unfamiliar situations they were not pre-trained and fine tuned for. This makes a difference to the creative intent of human beings. The relational space of AI technology for generating solutions is narrow and tends to follow a chain-like approach. This bears risks for unforseen situations and challenges demanding for contextualized new processes of generating solutions.

The movement towards human-centered AI (see Shneiderman 2022, 2020) aims at both, intelligent augmentation of human actors involved and a safty culture reflecting on the accountability at critical interfaces and reliable practices between the involved human and non-human entities. The vision is to "enhance human performance with systems that are reliable, safe, and trustworthy" (Fischer 2023, p. 1). These scholars claim the relational space that was often neglected when AI-based solutions became implemented but left critical interfaces without principles for reliable organizational practices (see Widder/Nafus 2023).

2.3. Relational practices as core of value creation – The service-dominant logic

It is exactly relational practices that is in the center of new business thinking and value creation. Under the lens of the service-dominant logic (SD logic, Vargo/Lusch 2004, 2008; Blaschke et al. 2019) scholars emphasize interaction, collaborative and co-creative practices as the core of relational value creation in business ecosystems (Vargo/Lusch 2016). There is a high potential for

DOI: https://doi.org/10.30844/wgab_2023_5

continuously interacting with customers, finding and further developing customer specific solutions for certain fields of established and new business (Coreynen et al. 2017). Even though the SD perspective has an almost twenty year record and was considered as crucial for manufacturing industries in Western economies (Zimmermann et al. 2021) the meaning of generating value from relational practices is not combined with specifications for the operational level in AI-based systems. Especially the crucial interfaces for generating solutions have been negelcted or underestimated (Thewes et al. 2022).

Paschen et al. (2021) analysed the human-AI-co-creation in sales marketing on empirical basis and describe that AI tools perform enabler and operator functions while human agents serve as experts, creators, conductors and reviewers. Galsgaard et al. (2022) make a conceptual outline for radiology and argue that a division of tasks and separation of expertise would be a constraint for technology implementation while a role concept of collective human-AI-expertise or sense of collective expertise would make systems work. Currently, there are scenarios for separated and integrated tasks or even role concepts in parallel but the meaning of relational practices on system level remains underexplored.

2.4. Facing the challenges of relational practices in semi-autonomous systems

In an ideal world there would be semi-autonomous systems allowing us to exploit the potential of AI while making use of human creative intent in order to benefit from system development with respect to single-loop and double-loop learning and thus enhance resilience (Evenseth et al. 2022). This would make systems robust and adaptable as they have a both-directional option for generating better solutions in terms of high quality, accuracy and safety (Shneiderman 2020; Widder/Nafus 2023). The prerequisite is that ML algorithms provide and sustain a space for human creative intent as valuable part of a co-created solution. Currently, semi-autonomous AI systems have limited agility in constructing experimental situations by which creative intent can be explored and unfolded. This can be illustrated by use fields of high relevance for society and a bright future.

Use case: semi-autonomous driving

The vision of autonomous driving is the most popular use field for reflecting (semi-)autonomous systems (e.g. Fischer 2018). It is an AI application with tremendous impact on societal level. The human side of the current discourse raises many questions including ethical issues if technology has to “decide” whether one persons' life goes over the other in face of unavoidable car accidents. Concepts for semi-autonomous assisted driving aim to keep the driver in the loop for critical situations. Thus, there might be a focus on human creative intent but the technical approach at the same time undermines it and raises certain questions as the outlines for the driver role are underdeveloped. How can a person who is not concentrated on driving and not trained by daily routines in driving a vehicle better regulate a critical situation spontaneously than the technology? What type of experience is necessary to perform this type of task? Is experience in non-assisted driving a prerequisite to manage these situations? But this would be only the case for the first generation of drivers adapting to semi-autonomous driving. Is the driving licence of the future rather a training in how autonomously driving systems work and how to interact with them in case of emergency – similar to the training for pilots. The questions show that the critical interfaces of (semi-)autonomous driving are not sufficiently reflected and that sophisticated human-AI-role concepts for making systems more reliable and safe are missing. Creative intent requires contextual knowledge and experience which is not considered in a meaningful manner as long as the system development follows the technological potential and treats the individual potential as residual factor. Dongol et al. (2020) highlight how interactions between environmental contextual information, autonomous systems, and regulation needs to be approached outside of existing models with contextually informed methods that go beyond hazard-focused and procedural compliance. The authors note however that challenges still remain with regard to finding appropriate test scenarios and that further work is necessary on the handover processes between AI and humans. In this space, reflective practices and particularly *reflection in action* – on the part of an autonomous system – may provide a theoretical context for re-assessing the human-AI interface. This would require another approach of technology development on the one hand side in experimental settings with the behavior of drivers in the center of analysis and a focus on reactions and reflections in order to find solutions for a corresponding supportive technology-based driver assistance. On the other hand ethnographic research observing the change of (driving) behavior and individual routines while being transported by a semi-autonomous vehicle needs to be analyzed in oder to understand new reflective practises instead of relying on former models of driving and drivers role that do not correspond with the new context factors.

Use case: manufacturing

Manufacturing is an advanced use field for integrating AI in production flows and work processes. The range goes from exoskeletons for physical health protection, eyeware for supporting operational tasks or digital assistance systems in production and supply chains in order to balance the mental and physical shape of an operating person but also for the overall condition monitoring at the critical interfaces (Romero et al 2016; Hinrichsen/Bendzioch 2019). The aim is to make systems more efficient and reliable for all stakeholders. This is complemented with training approaches for employees in how to handle and interact with the digital tool (e.g. Gorecky et al. 2014). Employees are considered as users and important actors of an implementation approach dedicated to the potential of the technology. The key challenge here is that the system design follows an engineering approach of standardising processes for production flows while considering the technology as a tool for regulating systems' need and compensating individual shortcomings

DOI: https://doi.org/10.30844/wgab_2023_5

in fatigue, disabilities etc. (Wilkens et al. 2021). The underlying process optimization reduces (unintendendly) the space for human creative intent for performative interaction for system regulation. It is not the individual who is considered to manage systems' weeknesses but the machine managing individual shortcomings. This can be acceptable for pre-planned chains following a good-dominant logic with standardizable transactions but can easily become a hazard for generating solutions in collaborative value creation networks.

Use case: Software development for industry applications

The software development of AI-tools to be applied in other user domains, e.g. manifacturing, healthcare or education, is decribed as a rather modularized process even within software companies depending on organizational internal checks and balances at critical interfaces of the development process including issues of data reliability and ethical guidelines that are brought to standards and checklists but however can easily be circumvented or at least not be treated in as deep sense as they are supposed to be. This is documented to be a challenge of organization and process chains in modularized development processes (Widder et al. 2021; Widder/Nafus 2023). In addition to this challenge, the domains are almost not involved in the development process. Data classification for software tool takes place without domain knowledge even though a contextualization is important for a meaningful use of data. This severe problem has especially been described for the heathcare sector (Thewes et al. 2002; Morrow et al. 2023). This raises questions with respect to the trustworthiness of AI applications as meaningful and reliable data classification needs the knowledge of user domains and not just of software developrs. This is almost not the case in daily practices. Tools dedicated to human-AI interaction miss human-human interaction in their process of development. The sense for interaction at critical interfaces is not a taken-for-granted routine in technology development. Interfaces are rather considered as a challenge of technological interoperability one aims to get rid of. As a consequense reflecting and designing the necessary space for creative intent and performative practices to make solutions better and reliable is not in the center of the process of software development. This is neither the case for organizational internal processes nor for processes between the developr and the user domain. The hidden ideal type stil tends to follow the vision of Industry 4.0 of autonomously interacting digital systems.

Use case: radiology

There are meanwhile visions for using AI for medical diagnoses in radiological images following an Industry 5.0 approach. Scholars describe radiologists as system regulators which make use of AI-applications in order to enhance the accuracy of their diagnoses and to better interact with other medical disciplines as an issue of system regulation (Dewey/Wilkens 2019). There is the idea of a collective expertise and related role development of radiologists (Galsgaard et al. 2022). The core idea is that AI supports the standard classification of images and provides a safety net for physicians who further validate their individual diagnosis and especially spend their time on critical cases. These concepts build on the creative intent of the professionals to which the technical system provides the space and therefore give a good example. Organizational process descriptions need to support this type of human-AI interaction. Currently, this is the struggeling point because concepts are projecting state-of-the-art interactions to the future and thus tend to neglect that standard operations have to be adapted when they are dedicated to high reliablity through human-AI system regulation and thus need to be build around human creative intent and its further development.

The first generation of radiologists working with AI-based image classification has a high profession and deep sense of interpreting images while making use of all human senses and contextual information. This allows them a continuous reflection on action and decision making while collaborating with AI. Future generations are trained on AI technology from the very beginning and thus need to develop an understanding when to trust and rely on the diagnosis proposed by AI and when to mis-trust it, how to develop the expertise to be able to mis-trust the technology as an important element of AI literacy and to be able to have the sense for another diagnosis. The space for creative intent has to be considered in role concepts. So far, it currently exists in the first user generation but tends to slowly disappear with future generations (similar to the future generations of drivers), especially if user domain knowledge is not sufficiently integrated in the software development process (see use case above). There is also a counterproductive side-effect from technology that has to be taken into consideration. AI speeds up the process of diagnosis - this is considered as an issue of enhancing productivity while integrationd AI – but this may unintentionally reduce the time for creative intent that cannot develop sufficiently in high speed settings as the load for individual reflection decreases with the amount of images to be reflected in a time unit. There are entire system dynamics that have to be taken into consideration in order to design hybrid systems based on human-AI-interaction in a manner that allows to reach the targets of high reliability in the long run.

Use case: Generative Pre-trained Transformer (GPT) applications in Higher Education

Natural Language Modeling is on the research agenda since 2010. Data scientists started to develop Large Language Foundation Models in 2017 and came up with the first public available Conversational Large Language Foundation Model in November, 30, 2022: Chat GPT III is the third generative pre-trained transformer (GPT) application offered to a broader group of users by OpenAI. The speed of global dissemination of the test version was higher than for any other technology before.

DOI: https://doi.org/10.30844/wgab_2023_5

Since then a generative AI-tool is in broad use in the societies all over the globe. The much more advanced and licenced professional version ChatGPT IV came out in March 2023. While ChatGPT III is based on 175 billion textual data, ChatGPT IV is based on 100 trillion textual data (Barthakur 2023).

A "time for class survey" from March 2023 (Student n=1,545, faculty n=1,692, admin n=205) (Bharadwaj et al. 2023) shows for the use field of higher education that in the first three months of the free version it was only 9% of faculty using the tool but 29% of students. Students prefere to use the tool for individual feedback and tutoring while writing assignments especially while brainstorming and structuring their essay but rather not for writing the whole text (Bharadwaj et al. 2023). All respondents of the survey are convinced that the tool improves individual learning strategies and learning outcomes. The licence price of the professional version is adapted to students' budget – at least in highly developd Western countries. Strategies for learning and generating solutions already changed among the future generation of global leaders and this is of impact for our future. The AI-based co-creation of learning processes and learning outcomes is on the daily agenda. What does this imply for keeping creative intent in the center of generating solutions? This is an open question demanding for further research. In the current state of development students tend to use and experiment carefully with a tool in order to generate better solutions. It seems to be part of the opacity that students show high consciousness for the possible but unknown weaknesses of the tool and thus intensify their reflective practises. This would be in line with the key targets of higher education, boostering the learning process towards human reflective practices (Barthakur et al. 2022). But this might change in future when students have higher trust in GPT outcomes and submit AI generated assignments which are in the next step evaluated by AI tools activated by faculty. In these constellations all involved parties have high proficiency in operating with AI-based digital tools to safe time for other tasks and interests but tend to loose creative intent in generating even better outcome in person. Keeping creative intent in the center of AI development is a challenge for the future of an intelligent humanity. The education system is the transformer for all other domains and thus needs to find solutions.

Providing systems with the capacity to continuously further develop and generate meaningful contextual interactions through hybrid human-technology reflective processes will be the core challenging issue for all institutional fields operating with AI.

3. Discussion

We outlined with the help of different use cases that the design of AI-based system solutions is often not dedicated to keep human creative intent continiously high. Even if there is an approach for the first adopters as decribes for healthcare and higher education, the design perspectives for future – not just on a technological basis, but also on an organizational basis – are not yet clear. In the use field of higher education there is the most explicit consciuousness for human reflective practices (Barthakur et al. 2022) but even this field is now challenged by generative AI applications with a high risk to act autonomously instead of interactive between individual learners and technology. There is a potential for humanity to steer into a bright future with semi-autonomous systems benefitting in their further development from individual learning and reflection and from machine learning mechanisms of data processing. But this is a narrow ridge that can easily reverse in an opposite direction. For the bright side system design principles and related organizational strategies are necessary, including technology development, organizational learning and individual learning & development. Corresponding methods have to complement each other in a consistent manner in order to enhance the reliability on system level. Coping with the risks of AI while exploiting the benefits is an issue of sociotechnical design and development.

Reflecting the state of the art descriptions of the use cases against the background of what it implies for keeping human creative intent in the loop is summarized in figure 1. Considering current design approaches in the light of individual competence development or space for reflective practices dedicated to creative intent leads to the conclusion:

(1) that is is especially the autonomy of systems that can disturb creative intent and negatively effects individual learning and development (see bottom of the figure),

(2) that most industry applications and taken-for-granted practices of buying and implementing AI-based solutions reduce the human side to rather uncritical users of AI tools with pure application skills (see middle of the figure) and

(3) that process outlines dedicated to high crerative intent are either missing or do only exist for early adapters or first generation users (top of the figure) while late adopters might only have application skills (see below).

DOI: https://doi.org/10.30844/wgab_2023_5

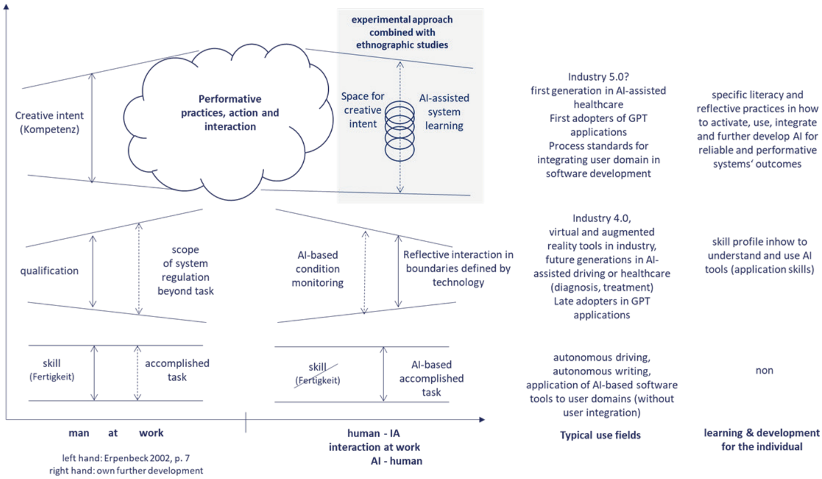

Figure 1: Design principles of human-AI interaction and their impact on performative practices through human creative intent

4. Outlook: An agenda for methodology development

Methodologies in AI development rather rely on optimization and outcome criteria instead of providing the necessary space and complementarity for benefitting from creative intent in order to reach higher targets in system level reliability. One of the key problems of modelling is that human creativity is not a fully generalisable system: there is no one flowchart of creativity, no singular set of processes. Instead, the interactions and networks that define creative responses to real-world situations occupy a layered, rhizomic configuration that resides between experience and experiment. This interplay needs to be addressed in future research approaches. The way in which this conceptual space of creative development operates shares much in common with organisational models such as SD logic. Yet, industrially, creative thought is often assumed to be part of a linear, goods-based logic – resulting in outcomes, design features, or traditional "creative" destinations such as branding, artwork or media applications. All of these situations represent product-based thought. Creativity itself is not an assembly line, or a set of prescribed processes. That doesn't mean that it functions without structures or relationships, and it is these which underpin decision making. SD logic and AI systems based on current machine learning or conversational large language models consisting of neural networks share similar architectures: it is not possible to find the "value" expressed directly in the structure of the network itself, it is created through the relational interaction of the component parts.

It is surprising therefore that SD logic-based workflows and AI systems which require contextual interaction are not generally associated in systems design processes. However, both methods depend on the similar relational information generation, on emergence and on pragmatic validity, to substantiate their activity.

Starting the other way round it is important to note that intent is subject to emergence. Creativity – as an idea – encompasses consideration of emergent behaviours and as such, has become polarised between views of creativity as straight forwards inspiration and views of creativity as fully-worked-through craft processes. In reality, neither is likely to be the sole driver for outcomes. Bringing creativity together with Rheinberger's (1997) ideas for experimental systems results in a consideration of ideas which are either directly expressed through design or those which arise through emergence. Current AI systems have difficulty in situating emergence, as it is of itself not a product of a technical rationality.

Established methods in co-design such as design thinking need to be revisited under the conditions of generative AI as development cyles remain to sequent like and to interconnected. Future research methods have to consider systems with meaningful contextual interactions with hybrid human-AI reflective processes. Crucial to forming a contemporary description of intent that is relevant for workplace semi-autonomous system design is the idea of reflective practice established by Schön (1983). Reflective practice is not an adjunct to professionalism, it is an integral part of being a professional practitioner. Reflective practices have been adapted and applied to high risk work environments and professional training scenarios in high reliability organizations with high responsibility in health and safety (Jordan, 2010). As Schön (1983) argues: this is of particular importance in areas with high situational awareness and mindeful enactment of routines where complex systems, creativity and professional practice interact. For Schön (1983), there are two key types of reflection: reflection as an activity in the abstract, and reflection in action. It is proposed here that reflection in action is critical to understanding the interaction between context and intent, and therefore, a critical process for underpinning semi-autonomous system design. Reflection in action has much in common with experimentation, in an artistic sense, rather than one inherited from the sciences with its clinical character of simulations. This is the concept where we propose to elaborate on in future research especially combined with ethnographic studies that are considered as most important for organizational settings with human-AI-interaction systems (Anthony et al. 2023; Widder/Nafus 2023). This allows to become part of an interactive community, e.g. of learners continiously using GPT applications and to observe how their way of asking questions, design solutions, trust and mis-trust a system develops over a longer time period.

"Experiment" in science is considered as a method with rigorous, perhaps linear, process of thought to deduct specific information about a specific problem, for example in trying to what the boundaries of tolerance are on an engineering design.

DOI: https://doi.org/10.30844/wgab_2023_5

By its very nature, the act of experimenting in a scientific context may be involved with working with a closed system: with deliberately limited use-cases, input conditions and a controlled environment. This is meaningful for testing tools (see bottom and middle of figure 1) but it is not a suitable approach for exploring the organic development of systems based on generative AI. Exploring the necessary space for creative intent requires to understand where are the roots of personal styles, new forms of interaction and related reflective practices with unexpected outcomes. This is the opposite to a science-based process validation. The core research question is how the relationships between human experience and experimental practice with AI-based applications emerge to form contextual knowledge that is essential to cope with unforseeable situations and high risks. The aim is to explore where are the critical spaces in which creative intent occurs and further develops, which situations and unstandardized context characteristics are necessary what is the time that needs to be reserved in technical supported systems for sustaining creative intent. Otherwise the augmentation potential of AI cannot be exploited. "Experimental designs" coping with these needs are living lab approaches that make contextualization to a design principle. This is why they are closely connected to ethnographic studies, e.g. in outlines of participatory studies on platforms. As far as the parameters for reflective practices have been explored an interacting and complementing technical system can be further developed. The challenge for future work is not to build AI after the human brain. The crucial question is how to develop AI that supports and sustains the creative intent that makes the difference in quality of life and quality of work in terms of safety, reliability and trust? Research facilities such as the research building ZESS for the engineering of smart product-service-systems provide such a research environment with living lab character (https://forschung.ruhr-uni-bochum.de/de/forschungszentrum-fuer-das-engineering-smarter-produkt-service-systeme-zess). And there are international counterparts, e.g. the institute and facilities for safe autonomy at the University of York, UK (https://www.york.ac.uk/safe-autonomy/facilities/). We invite to collaborate with us in the outlined inquiry for which engineering expertise is essential.

References

Anthony, C., Bechky, B. A., & Fayard, A.-L. (2023). "Collaborating" with AI: taking a system view to explore the future of work. Organization Science. https://doi.org/10.1287/orsc.2022.1651

Barthakur, A., Joksimovic, S., Kovanovic, V., Mello, R. F., Taylor, M., Richey, M., & Pardo, A. (2023). Understanding depth of reflective writing in workplace learning assessments using machine learning classification. IEEE Transactions on Learning Technologies, 15(5), 567–578. http://dx.doi.org/10.1109/TLT.2022.3162546

Bharadwaj, A. (2023). AI, generative AI and large language models [presentation]. Centre for Change and Complexity in Learning, University of South Australia. https://docs.google.com/presentation/d/1yL6sAW8vgam7K_1ASsYUq9gU1f5OP4XV/edit#slide=id.p5

Blaschke, M., Riss, U. V., Haki, K., & Aier, S. (2019). Design principles for digital value co-creation networks: a service-dominant logic perspective. Electronic Markets 29(3), 443–472. https://doi.org/10.1007/s12525-019-00356-9

Coreynen, W., Matthyssens, P., & van Bockhaven, W. (2017). Boosting servitization through digitization: pathways and dynamic resource configurations for manufacturers. Industrial Marketing Management, 60, 42–53. https://doi.org/10.1016/j.indmarman.2016.04.012

Dewey, J. (1934). Art as Experience. New York: Capricorn Books.

Dewey, M., & Wilkens, U. (2019). The bionic radiologist: avoiding blurry pictures and providing greater insights. Npj Digital Medicine, 2(65), 1-7. https://doi.org/10.1038/s41746-019-0142-9

Dongol, B., Bell, R., Habli, I., Lawford, M., Moore, P. J., & Saigol, Z. (2020). Panel discussion: regulation and ethics of robotics and autonomous systems. In A. Cavalcanti, B. Dongol, R. Hierons, J. Timmis, & J. Woodcock (Eds.), Software Engineering for Robotics (pp. 467–483). Cham: Springe. https://doi.org/10.1007/978-3-030-66494-7_14

Erpenbeck, J. (2002). Kompetenz und Performanz im Bild moderner Selbstorganisationstheorie. Dokumentation 4. BIBB-Fachkongress 2002.

Erpenbeck, J., von Rosenstiel, L., Grote, S., & Sauter, W. (Eds.) (2017). Handbuch Kompetenzmessung. Erkennen, verstehen und bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis (3rd ed.). Stuttgart: Schäffer-Poeschel.

Evenseth, L.L., Sydnes, M., & Gausdal, A. H. (2022). Building organizational resilience through organizational learning: a systematic review. Frontiers in Communication, 7, 837386. https://doi.org/10.3389/fcomm.2022.837386

Fagnant, D. J., & Kockelman, K. M. (2015). Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. Transportation Research Part A: Policy and Practice, 77, 167–181. https://doi.org/10.1016/j.tra.2015.04.003

Fischer, G. (2018). Exploring design trade-offs for quality of life in human-centered design. Interactions, 25(1), 26–33. https://doi.org/10.1145/3170706

DOI: https://doi.org/10.30844/wgab_2023_5

Fischer, G. (2023, June). A Research Framework Focused on AI and Humans instead of AI versus Humans. Proceedings of CoPDA2022 – Sixth International Workshop on Cultures of Participation in the Digital Age: AI for Humans or Humans for AI?.

Galsgaard, A., Doorschodt, T., Holten, A. L., Müller, F. C., Boesen, M. P., & Maas, M. (2022). Artificial intelligence and multidisciplinary team meetings; a communication challenge for radiologists' sense of agency and position as spider in a web? European Journal of Radiology, 155, 110231. https://doi.org/10.1016/j.ejrad.2022.110231

Gorecky, D., Schmitt, M., Loskyll, M., & Zühle, D. (2014, July 27–30). Human-machine-interaction in the industry 4.0 era [conference paper]. 12th IEEE International Conference on Industrial Informatics (INDIN), Porto Algere.. https://doi.org/10.1109/INDIN.2014.6945523

Haenlein, M., & Kaplan, A. (2019). A brief history of Artificial Intelligence: on the past, present, and future of Artificial Intelligence. California Management Review, 61(4), 5–14. https://doi.org/10.1177/0008125619864925

Hashimoto, D. A., Rosman, G., Rus, D. & Meireles, O. R. (2018). Artificial intelligence in surgery: promises and perils. Annals of Surgery, 268(1), 70–76. https://doi.org/10.1097/sla.0000000000002693

Heyse, V., & Erpenbeck, J. (2007). Kompetenzmanagement: Methoden, Vorgehen, KODE® und KODE®X im Praxistest. Münster: Waxmann.

Hinrichsen, S., & Bendzioch, S. (2019). How digital assistance aystems improve work productivity in assembly. In I. L. Nunes (Eds.), Advances in Human Factors and Systems Interaction (pp. 332–342). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-94334-3_33

Jepsen, S. C., Mørk, T. I., Hviid, J., & Wormet, T. (2020, December 14-17). A Pilot Study of Industry 4.0 Asset Interoperability Challenges in an Industry 4.0 Laboratory [Conference paper]. 2020 International Conference on Industrial Engineering and Engineering Management. Singapore.

Jordan, S. (2010). Learning to be surprised: how to foster reflective practice in a high-reliability context. Management Learning, 41(4), 391–413. https://doi.org/10.1177/1350507609357388

Kaartemo, V., & Helkkula, A. (2018). A systematic review of Artificial Intelligence and robots in value co-creation: current status and future research avenues. Journal of Creating Value, 4(2), 211–228. http://dx.doi.org/10.1177/2394964318805625

Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. Business & Information Systems Engineering, 6, 239–242. https://doi.org/10.1007/s12599-014-0334-4

Launchbury, J. (2017, February 15). A DARPA perspective on Artificial Intelligence. Retrieved from https://machinelearning.technicacuriosa.com/2017/03/19/a-darpa-perspective-on-artificial-intelligence/

Maeda, J. (2002). Design education in the post-digital age. Design Management Journal (Former Series), 13(3), 39–45.

Morrow, E., Zidaru, T., Ross, F., Mason, C., Patel, K. D., Ream, M. & Stockley, R. (2023). Artificial intelligence technologies and compassion in healthcare: A systematic scoping review. Frontiers in Psychology, 13. https://doi.org/10.3389/fpsyg.2022.971044

Nahavandi, S. (2019). Industry 5.0 – A human-centric solution. Sustainability, 11(16), 4371. https://doi.org/10.3390/su11164371

Neblo, M. A. (2007). Philosophical Psychology with Political Intent. In G. E. Marcus, W. Russel Neuman & M. MacKuen (Eds.), The Affect Effect: Dynamics of emotion in political thinking and behavior (pp. 25–47). Chicago: University of Chicago Press.

Paschen, J., Paschen, U., Pala, E., & Kietzmann, J. (2021). Artificial intelligence (AI) and value co-creation in B2B sales: Activities, actors and resources. Australasian Marketing Journal, 29(3), 243–251. https://doi.org/10.1016/j.ausmj.2020.06.004

Rheinberger, H.-J. (1997). Experimental complexity in biology: Some epistemological and historical remarks. Philosophy of Science, 64(4), 245–54. https://www.jstor.org/stable/188407

Romero, D.,Stahre, J.,Wuest, T., Noran, O.,Bernus, P., Fast-Berglund, Å., & Gorecky, D. (2016, October 29-31). Towards an operator 4.0 typology: a human-centric perspective on the fourth industrial revolution technologies [Conference paper]. Proceedings of the international conference on computers and industrial engineering (CIE46). Tianjin.Rogala, J., Bajno, B., & Wróbel, A. (2020). A hidden message: Decoding artistic intent. PsyCh Journal, 9(4), 507–512. https://doi.org/10.1002/pchj.374

Russel, S., & Norvig, P. (2021). Artificial Intelligence, Global Edition. A Modern Approach (4th ed.). Edinburgh: Pearson Education.

Schön, D. A. (1983). The Reflective Practitioner: How Professionals Think in Action. New York: Basic Books.

Shin, R. (2023, June 8). Sam Altman, the man behind ChatGPT, is increasingly alarmed about what he unleashed. Here are 15 quotes charting his descent into sleepless panic. https://fortune.com/2023/06/08/sam-altman-openai-chatgpt-worries-15-quotes/

Shneiderman, B. (2020). Human-centered Artificial Intelligence: Reliable, safe & trustworthy. International Journal of Human–Computer Interaction, 36(6), 495–504. https://doi.org/10.1080/10447318.2020.1741118

Shneiderman, B. (2022). Human-Centered AI. Oxford: Oxford University Press.

Thewes, F., Langholf, V., Meske, C., & Wilkens, U. (2022). Towards a process model for cross-domain AI development – Insights from neuroradiological imaging. In Tagungsband GfA Frühjahrskongress 2022 (Eds.), Technologie und Bildung in hybriden Arbeitswelten, Sankt Augustin, Magdeburg.

Van Rijmenam, M., & Logue, D. (2020). Revising the `science of the organisation`: theorising AI agency and actorhood. Organization & Management, 23(1), 127–144. https://doi.org/10.1080/14479338.2020.1816833

Vargo, S. L. & Lusch, R. F. (2004). Evolving to a new dominant logic for marketing. Journal of Marketing, 68, 1–17.

Vargo, S.L. and Lusch, R.F. (2016) Institutions and axioms: An extension and update of service-dominant logic. Journal of the Academy of Marketing Science, 44, 5–23. https://doi.org/10.1007/s11747-015-0456-3

Widder, D. G., Dabbish, L., Herbsleb, J. D., Holloway, A., & Davidoff, S. (2021, May 8-13). Trust in collaborative automation in high stakes software engineering work: A case study at NASA [Conference paper]. Conference on Human Factors in Computing Systems. Yokohama.

Widder, D. G., & Nafus, D. (2023). Dislocated accountabilities in the "AI supply chain": Modularity and developers` notions of responsibility. Big Data & Society, 10(1), 1–12. https://doi.org/10.1177/20539517231177620

Wilkens, U. (2020). Artificial intelligence in the workplace – A double-edged sword. International Journal of Information and Learning Technology, 37(5), 253–265. https://doi.org/10.1108/IJILT-02-2020-0022

Wilkens, U., Keller, H., & Schmette, M. (2006). Wirkungsbeziehungen zwischen Ebenen individueller und kollektiver Kompetenz. Theoriezugänge und Modellbildung. In G. Schreyögg & P. Conrad (Eds.), Management von Kompetenz (pp. 121–161). Wiesbaden: Gabler.

Wilkens, U., Langholf, V., Ontrup, G., & Kluge, A. (2021). Towards a maturity model of human-centered AI – A reference for AI implementation at the workplace. In W. Sihn, & S. Schlund (Eds.), Competence development and learning assistance systems for the data-driven future (pp. 179–197), Berlin: GITO-Verlag.

Wilkens, U., & Sprafke, N. (2019). Micro-variables of dynamic capabilities and how they come into effect – Exploring firm-specificity and cross-firm commonalities. Management International, 23(4), 30–49. http://dx.doi.org/10.7202/1066068ar

Xu, W. (2019) Toward human-centered AI: a perspective from human-computer interactions. Interactions, 26(4), 42–46. https://doi.org/10.1145/3328485

Zimmermann, A., Schmidt, R., Sandkuhl, K., & Masuda, Y. (2021). Architecting Intelligent digital systems and services. In A. Zimmermann, R. Howlett, L. Jain (Eds.), Human Centered Intelligent Systems (pp. 127–137). Springer. https://doi.org/10.1007/978-981-15-5784-2_11