

**Navigating the intersection of developers, domain experts and disruptive AI:
An identity work perspective**

Lara Carminati¹, Jeroen Meijerink¹, Nadine Voß¹

¹University of Twente, Faculty of Behavioural, Management and Social Sciences

Corresponding author:

Lara Carminati: l.carminati@utwente.nl

Navigating the intersection of developers, domain experts and disruptive AI:

An identity work perspective

Abstract: This developmental paper reports a work-in-progress study that aims at exploring how AI, seen simultaneously as the product of, and a threat to, an interprofessional collaboration between developers and domain experts, can influence the identity work of its creators. Through semi-structured interviews, we plan to conduct an inductive qualitative study among developers and domain specialists designing and implementing AI in their daily practice and analyse the data with grounded theory approaches. The purpose of this paper is to contribute to identity work and AI literatures by understanding how AI, as a unique, ‘quasi-social’ agent and source of identity threat, can disruptively interact with developers and domain experts and induce their identity work in a collaborative circle of creation.

Word count: 1,874

1. Introduction

Over the past decades, technologies have advanced at an exponential rate leaving almost no parts of humans' lives untouched (Byrnjolfsson & Macafee, 2014). Among these technologies, artificial intelligence (AI) is considered one of the most disruptive, and one that is likely to radically change the world of work. Defined as a machine that can perform functions attributed to human intelligence (Byrnjolfsson & Macafee, 2014), AI is able to produce knowledge based on data input by developers and independent from domain experts (van den Broek et al., 2021). Due to its independence, AI is claimed to even surpass domain experts' insights by producing knowledge free of human bias, discovering *new* information, and working at an unattainable speed (van den Broek et al., 2021). However, what seems like AI's greatest advantage, i.e., an objective, up-to-date and fast-accessible knowledge, can turn into its greatest disadvantage, since the produced knowledge can lack relevance to its users, i.e., the domain experts (Amershi et al., 2014; Holzinger, 2016; van den Broek et al., 2021). Hence, designing a useful AI tool for complex knowledge work requires interprofessional collaborations on behalf of both software developers and domain specialists and/or experts (González-Gonzalo et al., 2022; van den Broek et al., 2021).

The dynamics of interprofessional collaborations have been well studied in relation to individuals' identities and identity work (Ahuja, 2022; Barbour & James, 2015; Caldwell et al., 2017). Individuals possess multiple identities that reflect various social aspects of their lives (Ramarajan, 2014) and provide them with meanings and expectations that drive their behaviours (Brown, 2017; Ashforth & Schinoff, 2016). Hence, an individual's identity has been defined as a person's answer to the question 'who am I?' with the aim of maintaining a coherent sense of self (Alvesson et al., 2008). To maintain such a unique and coherent sense of self, individuals engage in identity work, the process of constantly making changes, revising or trying to strengthen their identities (Sveningsson & Alvesson, 2003). Identity work is particularly important during interprofessional collaborations since such collaborations can draw the identities and values of the different professions together. Although some research has shown that such collaborations could also lead to identity enhancement (Ramarajan et al., 2017), other studies have underlined how working with other professionals may trigger some conflicts and tensions (Brown, 2017; Horton et al., 2014). These tensions may result from discrepant professional values *within* the same identity or *between* different identities (Carminati & Héliot, 2021; 2022), or even from identity threat, a feeling of "potential harm to the value, meanings, or enactment of an identity" (Petriglieri, 2011, p. 644).

The literature acknowledges three main sources of identity threats that can induce identity work: the individual, the social world and the material world (Petriglieri, 2011). Collaborations with other professionals could be considered as threats from the social world since these collaborations take the form of interpersonal interactions in which professional identity values and beliefs from different social groups may clash. If there is inconsistency between individuals' actions and the meanings that the individual associates with a specific identity, a threat could emerge (Anteby, 2008). Nevertheless, when software developers and domain experts collaborate on a professional level to design AI, then the impact of AI implementation itself could become another threat to, mostly, domain experts. Software developers inscribe their values and norms into AI which, as a source of power (Orlikowski, 1992), enables them to impose their (alternative) worldview onto domain experts, creating potential identity conflicts. In a sort of vicious, iterative cycle, who helped co-developing AI could then be threatened by AI itself (Jussupow et al., 2022), especially if AI outcomes start questioning the professional judgment, values and principles of the domain experts.

However, not only we still know little of this 'AI-driven identity conflict', but it is also difficult to classify AI into one of the three sources of threat mentioned above due to three

main reasons. Firstly, AI may appear as a quasi-social third actor to professionals due to its self-learning capacities and autonomous deployment (Selenko et al., 2022). It is thus neither an object from the material world –since it has the possibility to interact with individuals–, nor can it be fully treated as another person. Secondly, the complex nature of its underlying algorithms may cause AI to be perceived as a “black box” especially by domain experts and/or specialists. Not being able to understand how the AI generates knowledge or verify its decisions, may lead to uncertainty and mistrust, and thus amplify the experienced identity threat by domain experts, and potentially, software developers too (Selenko et al., 2022; Strich et al., 2021). Lastly, since AI has the potential to substitute human work processes when fully developed (Strich et al., 2021), the identity threat may be continuously reinforced through the iterative cycle and nature of AI design. The more frequently an experience is encountered by an individual, the more likely identity threat is to occur, urging professionals to perform identity work to revise and strengthen their professional identity (Petriglieri, 2011).

And yet, perhaps due to the “black box” and threats surrounding AI design and implementation, it is still unclear how AI interplays with the identity values of professionals and influences their identity work. Existing studies have shown that AI may be perceived as an identity threat (Jussupow et al., 2022), but have mostly overlooked how professionals respond to this threat in form of identity work. Furthermore, studies have centered around domain experts in the role of AI users, but not also in the role of co-designers (Mirbabaie et al., 2021; Strich et al., 2021). However, scholars have advocated shifting the focus away from purely the use of technology more towards the ideology and values of the professionals involved in its design and creation (Bailey & Barley, 2019). Understanding how the values of developers shape AI design may help to overcome common issues associated with technology, such as ethical problems or user resistance (Bailey & Barley, 2019), especially since in a broader definition design not only involves creating and shaping complex systems, but also their implementation and ongoing adjustments (Freeman & Hart, 2004). Therefore, exploring how AI can be seen as the product of, and a threat to, an interprofessional collaboration between developers and domain experts and influence their identity work remains unique. Hence, since how professionals forge AI through their own process of identity work has been rather overlooked, this developmental paper aims at answering the following research question: **How does designing (and implementing) AI influence developers’ and domain experts’ identity work in their interprofessional collaborations?**

By answering this research question, this paper aims to contribute to the literature on identity work and AI in two ways. Firstly, by approaching AI through the lens of a collaborative design process, we shed light on the identity work that both developers and domain experts need to enact when creating AI. Secondly, by considering AI as a unique quasi-social agent, we extend our current knowledge on the theorizing on identity threat by illuminating how AI implementation may induce further identity work in those very same professionals who designed it.

2. Methods

2.1. Research design

For this paper we aim at conducting an inductive, qualitative study using grounded theory approaches (Strauss & Corbin, 1990) for two reasons: (1) given that the purpose of grounded theory approaches is to extend current, or develop new, theoretical insights into how individuals understand and make sense of reality subjectively, they are particularly appropriate for studying deeply subjective phenomena like identity work and identity threat (Suddaby, 2006); (2) grounded theory approaches are often well-suited when the focus of the research is to unpack processes and patterns in human behaviours related to the experience

under examination (Suddaby, 2006). Through an interpretive discursive lens (Miles & Huberman, 1994), our purpose is thus to produce a rich account of participants' perspectives by interpreting and giving meaning to their voices and narratives in order to extend current theories on identity work through the emergence of novel themes (Strauss & Corbin, 1990).

2.2. Sampling & Data Collection

To investigate how AI influences the identity work of professionals, we plan to select a sample of software developers and HR specialists and/or domain experts that design AI for human resource management (HRM). Since AI is gaining increasing importance in HRM practices and is already used for tasks such as recruitment and selection of personnel, this choice was deemed particularly suitable (Charlwood & Guenole, 2022). Moreover, research has shown that AI-based HRM requires a tight interplay and collaboration between software developers and HR specialists, with each group acting on different sets of values, norms and worldviews (Van den Broek et al., 2022). This implies that AI-driven identity conflicts may arise in these inter-professional collaborations.

2.3. Research Instruments

We aim at interviewing roughly 30 participants between developers and HR specialists/domain experts during the period April-June 2023. The interviews are semi-structured and conducted whenever possible face to face, but we are open to have video conferencing interviews if necessary. We chose semi-structured interviews since, whilst facilitating the identification and comparison of the emerging themes through a standard set of questions (Miles & Huberman, 1994), they also allow the interviewer to tailor the questions according to the interviewee's answers (Silverman & Marvasti, 2008). This is deemed particularly important to reach a greater depth of content and to uncover people's sense of self and identity (Miles & Huberman, 1994). Interviews are planned to last about one hour, to be audio recorded and transcribed verbatim, concealing any personal information or details from participants. Ethical approval to conduct the research is obtained from the University Ethics Committee.

2.4. Data Analysis

We plan to analyse the data in an iterative way, constantly comparing the data and the previously coded texts with new coded sentences and the emerging structure of themes (Strauss & Corbin, 1990; Miles & Huberman, 1994). Through discarding some schemes, merging similar codes, and adapting others, this process allows us to gather stronger evidence in our data any time we see a new conceptual category emerging. In this phase, it is important that we engage in deep and long discussions regarding the potential interpretations and structure of the data to establish a strong inter-rater reliability. Furthermore, we will make sure to rely strictly on the language used by the participants to frame issues and concepts in our findings to maintain a connection with the raw data. When no new codes emerge, this will give us confidence that we have reached saturation, the point at which further data will yield redundant insights (Strauss & Corbin, 1990).

3. Results

This paper is a work in progress as part of a Master thesis project. At the point of the submission, data collection has not yet started, but it is likely to happen between April and May 2023 with the data analysis planned between June and July 2023. Although we acknowledge the possibility of further refinements in the theoretical background, the focus of the next steps is on gathering our corpus of data and understanding it. This is likely to allow

us –had the paper to be accepted– to have some preliminary results to present during the BAM Conference round-table discussions through which we hope to get valuable feedback for the development of the full paper.

References

- Ahuja, S. (2022). Professional Identity Threats in Interprofessional Collaborations: A Case of Architects in Professional Service Firms. *Journal of Management Studies*, 60(2), 428–453. <https://doi.org/10.1111/JOMS.12847>
- Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014). Power to the People: The Role of Humans in Interactive Machine Learning. *AI Magazine*, 35(4), 105–120. <https://doi.org/10.1609/AIMAG.V35I4.2513>
- Bailey, D. E., & Barley, S. R. (2019). Beyond design and use: How scholars should study intelligent technologies. *Information and Organization*, 30(2), 100286.
- Barbour, J. B., & James, E. P. (2015). Collaboration for compliance: Identity tensions in the interorganizational and interdisciplinary regulation of a toxic waste storage facility. *Journal of Applied Communication Research*, 43(4), 363–384. <https://doi.org/10.1080/00909882.2015.1083601>
- Beck, A. H., Sangoi, A. R., Leung, S., Marinelli, R. J., Nielsen, T. O., van de Vijver, M. J., West, R. B., van de Rijn, M., & Koller, D. (2011). Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Science Translational Medicine*, 3(108), 108–113. <https://doi.org/10.1126/SCITRANSLMED.3002564>
- Brown, A. D. (2017). Identity Work and Organizational Identification. *International Journal of Management Reviews*, 19(3), 296–317. <https://doi.org/10.1111/IJMR.12152>
- Brown, A. D., & Coupland, C. (2015). Identity threats, identity work and elite professionals. *Organization Studies*, 36(10), 1315–1336. <https://doi.org/10.1177/0170840615593594>
- Byrnjolfsson, E., & Macafee, A. (2014). *The second machine age: Work, progress and prosperity in the time of brilliant technologies*. W.W. Norton & Company.
- Caldwell, N. D., Roehrich, J. K., & George, G. (2017). Social Value Creation and Relational Coordination in Public-Private Collaborations. *Journal of Management Studies*, 54(6), 906–928. <https://doi.org/10.1111/JOMS.12268>
- Carminati, L., & Gao Héliot, Y. (2021). Multilevel dynamics of moral identity conflict: professional and personal values in ethically-charged situations. *Ethics & Behavior*, 33(1), 37–54. <https://doi.org/10.1080/10508422.2021.2004891>
- Carminati, L., & Héliot, Y. G. (2022). Between Multiple Identities and Values: Professionals' Identity Conflicts in Ethically Charged Situations. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.813835>
- Charlwood, A., & Guenole, N. (2022). Can HR adapt to the paradoxes of artificial intelligence? *Human Resource Management Journal*, 32(4), 729–742. <https://doi.org/10.1111/1748-8583.12433>
- Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, 13(1), 3–21. <https://doi.org/10.1007/BF00988593/METRICS>
- Dreyfus, H. L. (2007). Why Heideggerian AI failed and how fixing it would require making it more Heideggerian. *Philosophical Psychology*, 20(2), 247–268. <https://doi.org/10.1080/09515080701239510>

- Freeman, P., & Hart, D. (2004). A science of design for software-intensive systems. *Communications of the ACM*, 47(8), 19–21. <https://doi.org/10.1145/1012037.1012054>
- González-Gonzalo, C., Thee, E. F., Klaver, C. C. W., Lee, A. Y., Schlingemann, R. O., Tufail, A., Verbraak, F., & Sánchez, C. I. (2022). Trustworthy AI: Closing the gap between development and integration of AI systems in ophthalmic practice. *Progress in Retinal and Eye Research*, 90, 101034. <https://doi.org/10.1016/J.PRETEYERES.2021.101034>
- Hogg, M. A. T. D. J. W. K. M. (1995). A tale of two theories: A critical comparison of Identity Theory with Social Identity Theory. *Social Psychology Quarterly*, 58(4), 255–269.
- Holzinger, A. (2016). Interactive machine learning for health informatics: when do we need the human-in-the-loop? *Brain Informatics*, 3(2), 119–131. <https://doi.org/10.1007/S40708-016-0042-6/FIGURES/1>
- Horton, K. E., Bayerl, P. S., & Jacobs, G. (2014). Identity conflicts at work: An integrative framework. *Journal of Organizational Behavior*, 35(1), 6–22. <https://doi.org/10.1002/JOB.1893>
- Jussupow, E., Spohrer, K., & Heinzl, A. (2022). Identity threats as a reason for resistance to artificial intelligence: survey study with medical students and professionals. *JMIR Formative Research*, 6(3), e28750. <https://doi.org/10.2196/28750>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook* (2nd ed.). Sage Publications, Inc. <https://doi.org/10.2307/1163741>
- Mirbabaie, M., Brünker, F., Möllmann, N. R., & Stieglitz, S. (2022). The rise of artificial intelligence—understanding the AI identity threat at the workplace. *Electronic Markets*, 32(1), 73–99. <https://doi.org/10.1007/S12525-021-00496-X>
- Nilsson, N. J. (1971). *Problem-solving methods in artificial intelligence*. McGraw-Hill.
- Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations. *Organization Science*, 3(3), 398–427.
- Petriglieri, J. L. (2011). Under Threat: Responses to and the Consequences of Threats to Individuals' Identities. <https://doi.org/10.5465/Amr.2009.0087>, 36(4), 641–662. <https://doi.org/10.5465/AMR.2009.0087>
- Pullon, S., Morgan, S., Macdonald, L., McKinlay, E., & Gray, B. (2016). Observation of interprofessional collaboration in primary care practice: A multiple case study. *Journal of Interprofessional Care*, 30(6), 787–794. <https://doi.org/10.1080/13561820.2016.1220929>
- Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, 46(1), 192–210. <https://doi.org/10.5465/AMR.2018.0072>
- Salovaara, A., Lyytinen, K., & Penttinen, E. (2019). High reliability in digital organizing: Mindlessness, the frame problem, and digital operations. *MIS Quarterly: Management Information Systems*, 43(2), 555–578. <https://doi.org/10.25300/MISQ/2019/14577>
- Selenko, E., Banks, S., Shoss, M., Warburton, J., & Restubog, S. L. D. (2022). Artificial Intelligence and the Future of Work: A Functional-Identity Perspective. *Current Directions in Psychological Science*, 31(3), 272–279. https://doi.org/10.1177/09637214221091823/ASSET/IMAGES/LARGE/10.1177_09637214221091823-FIG1.JPEG
- Serpe, R. T. (1987). Stability and Change in Self: A Structural Symbolic Interactionist Explanation. *Social Psychology Quarterly*, 50(1), 44–55.

- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press.
- Strich, F., Mayer, A. S., & Fiedler, M. (2021). What Do I Do in a World of Artificial Intelligence? Investigating the Impact of Substitutive Decision-Making AI Systems on Employees' Professional Role Identity. *Journal of the Association for Information Systems*, 22(2), 304–324. <https://doi.org/10.17705/1JAIS.00663>
- Suddaby, R. (2006). From the editors: What grounded theory is not. *Academy of Management Journal*, 49(4), 633–642. <https://doi.org/10.5465/AMJ.2006.22083020>
- Sveningsson, S., & Alvesson, M. (2003). Managing managerial identities: Organizational fragmentation, discourse and identity struggle. *Human Relations*, 56(10), 1163–1193.
- van den Broek, E., Sergeeva, A., & Huysman, M. (2021). When the machine meets the expert: An ethnography of developing AI for hiring. *MIS Quarterly*, 45(3), 1557–1580. <https://doi.org/10.25300/MISQ/2021/16559>