Perceived Authenticity in AI Afterlife Agents: A Dimensional Conceptualization

Mirella Moser and Jan Steiner
Zurich University of Applied Sciences, Institute of Information Systems, Winterthur, Switzerland

mosl@zhaw.ch
steinja8@students.zhaw.ch

Abstract: Advances in generative AI have enabled AI Afterlife agents that simulate deceased or absent persons through interactive systems, raising the question of what makes such representations appear authentic to close others. This paper conceptualizes perceived authenticity as the perceived fit between a digital representation and the remembered person and derives seven theoretically grounded dimensions: *Traits*, *Facts*, *Behavior*, *Decision Logic*, *Values*, *Boundaries*, and *Style*. The dimensions are implemented in a scenario-based prototype with an assessment and simulation phase and explored in a within-subjects study with close others of represented persons. Baseline scenarios were rated highly, while targeted mismatches in *Facts* and *Style* reduced perceived authenticity. *Facts* emerged as particularly relevant, suggesting that factual correctness may function as a basic hygiene factor that becomes salient when violated. The findings support the usefulness of a dimensional approach for designing and evaluating AI Afterlife agents, while also indicating that users may rely on fewer and more immediately recognizable cues when judging authenticity. Given the small exploratory sample and the manipulation of only two dimensions, the full seven-dimensional framework requires further empirical examination.

Keywords: AI Afterlife, Digital afterlife industry, Perceived authenticity, Digital legacy, Human-AI interaction, AI agents

1. Introduction

Advances in generative AI have accelerated the growth of the Digital Afterlife Industry (DAI), enabling services that simulate individuals after death through post-mortem re-creation (Hollanek & Nowaczyk-Basińska, 2024; Öhman & Floridi, 2018). Current debates on such AI Afterlife agents are often polarized. Some accounts frame these technologies as promising tools for remembrance and grief support (Lei et al., 2025; Lu et al., 2026). Others describe them as digital necromancy, manipulation, or a violation of dignity (Figueroa-Torres, 2024; Hollanek & Nowaczyk-Basińska, 2024; Jiménez-Alonso & Brescó de Luna, 2023). This debate is intensified by the difference between static digital traces and generative AI systems (Kidd & McAvoy, 2025). Photos, recordings, archived messages, and social media profiles have an evidential relation to the past: they preserve something the person actually said, did, or left behind. AI Afterlife agents, by contrast, generate new content in the present. They may therefore feel more lifelike, while their outputs are less securely tied to what the deceased actually said, intended, or endorsed (Hollanek & Nowaczyk-Basińska, 2024; Zhang et al., 2025).

Authenticity is becoming a central analytic category in this context because the decisive question concerns the kind of resemblance users experience when interacting with an AI Afterlife agent. Authenticity appears to shape whether such systems are experienced as comforting, misleading, or dependency-forming. Realism may increase emotional benefit (Sousa et al., 2025), but it can also increase the risk that users treat the system as more than a simulation (Spitale & Germani, 2025). More broadly, perceived authenticity has been shown to predict attachment and loyalty in human-AI relations (Pandey & Rai, 2024). Ethics literature treats authenticity as a moral issue because inauthentic simulations may distort memory, dignity, and relational trust (Fu et al., 2025). HCI research similarly identifies identity consistency and content authenticity as design problems central to the acceptance of AI Afterlife agents (Lei et al., 2025; Zhang et al., 2025).

However, authenticity remains undertheorized in this domain. It is often invoked to describe realism, identity consistency, emotional plausibility, or ethical acceptability (Hollanek & Nowaczyk-Basińska, 2024; Lei et al., 2025; Zhang et al., 2025), yet these dimensions are rarely distinguished. This ambiguity is consequential because in our eyes, an AI Afterlife agent can be perceived as authentic for different reasons: for example because it is grounded in personal data, because it behaves in recognizable ways, because it aligns with the represented person's values, or because it is experienced as relationally present by users.

2. Objective and Research Questions

This paper asks how authenticity can be conceptualized for AI Afterlife agents. More specifically, it examines which dimensions of authenticity are relevant when a system generates posthumous interaction in the name of a specific person. By distinguishing these dimensions, the paper contributes a framework for assessing when such agents can be understood as authentically connected to the person they represent. The central research questions are therefore:

- Which dimensions are theoretically relevant for conceptualizing perceived authenticity in AI Afterlife agents?
- How can these dimensions be operationalized in a prototype that simulates a specific reference person?
- Which authenticity dimensions are most relevant for close others when evaluating whether an AI Afterlife agent feels authentic or inauthentic?

3. Methodology

This research is grounded in the Design Science Research (DSR) framework (Hevner et al., 2004) and follows the Design Science Research Methodology proposed by Peffers et al. (2007). The research problem and objective were defined in advance: the study seeks to understand which dimensions make an AI Afterlife agent appear authentic to close others, and how such dimensions can be translated into an interactive prototype. The research was structured into three main phases:

First, the theoretical foundations of perceived authenticity were developed. Based on literature on personality research, digital legacy, AI Afterlife, authenticity, values, boundary management and explainable AI, relevant dimensions of authenticity were identified and systematized. The aim of this phase was to clarify what it means for a digital simulation to represent a particular person in a recognizable and fitting way. The resulting model conceptualizes authenticity as perceived fit between the AI Afterlife agent and the remembered person.

Second, the research artifact was designed and implemented. The prototype consists of two connected components. The Assessment Bot was developed to collect structured information about the (living) reference person along the seven authenticity dimensions. This information was then condensed into a reference profile. The Simulation Bot used this profile to generate scenario-based interactions for participants. The prototype was implemented with the PROMISE Framework (Wu et al., 2026/2023), which allowed the interaction logic to be modelled through states and transitions.

Third, the prototype was evaluated through a mixed-methods study with close others of the represented person. The empirical evaluation used a scenario-based within-subjects design. Participants interacted with the Simulation Bot while the living reference person was absent. They assessed several responses of the AI Afterlife agent using a standardized questionnaire. Quantitative data were analyzed descriptively and exploratively. Open-ended questionnaire responses were used as qualitative complements to interpret how participants justified their authenticity judgments.

4. Related Work

The development of AI Afterlife agents requires a differentiated understanding of how users assess the authenticity of a digital representation. Authenticity is generally understood as a multidimensional construct. In psychology, it is commonly associated with self-congruence, referring to the extent to which a person's behavior aligns with internal attitudes, values, and beliefs (Jongman-Sereno & Leary, 2019; Nunes et al., 2021). Kernis and Goldman (2006) similarly describe authenticity as a state in which inner experience, self-understanding, behavior, and relational conduct are coherently connected to the self. Consumer research conceptualizes authenticity as a formative construct comprising dimensions such as accuracy, integrity, and originality (Nunes et al., 2021). Moral philosophy extends this perspective through the concept of second-person authenticity, in which authenticity is assessed by another person within a relationship rather than by the self (Battisti, 2025).

This relational perspective is particularly relevant for AI Afterlife agents. In this context, authenticity is assessed by relatives or close others who compare the digital representation with their memories and expectations of the reference person. The relevant question is therefore whether the agent appears to correspond to a specific remembered individual. Such judgements may be sensitive to relatively small deviations. Human-like representations can cause discomfort when they appear almost, but not fully, similar to a real person (Mori et al., 2012).

In AI-mediated interaction, authenticity must also be distinguished from related concepts. Anthropomorphism refers to the attribution of human characteristics to non-human entities and may facilitate social interaction with AI systems (Li & Sung, 2021; Pandey & Rai, 2024). Trust is typically based on perceptions of competence, reliability, or benevolence (Mayer et al., 1995; Sonlu et al., 2021). Realism and factual accuracy concern the plausibility or correctness of generated outputs, but do not necessarily ensure that an agent reproduces a person's recognizable language, values, reasoning, or behavioral patterns (Riccardo et al., 2026). An agent may

therefore appear human-like, trustworthy, realistic, or factually correct without being perceived as an authentic representation of a specific person.

Research on persona consistency and identity representation provides a related foundation. Personas are commonly defined through combinations of identity-related facts, language patterns, and interaction styles (Li et al., 2016). However, conversational agents frequently encounter the out-of-predefined-persona problem, in which they generate unsupported or inconsistent characteristics when interactions exceed the information contained in the predefined profile (Liu et al., 2022). More recent approaches incorporate multimodal personality expression and broader criteria of recognizability across situations (Sonlu et al., 2021). Riccardo et al. (2026), for example, emphasize whether an agent remains an interactionally plausible representation across dimensions such as value coherence and situational response patterns.

In the context of digital afterlife systems, research has shifted from static digital archives toward interactive deadbots that support continuing bonds through the preservation and activation of digital legacy (Hollanek & Nowaczyk-Basińska, 2024; Lei et al., 2025). This transition changes the authenticity problem. Static materials preserve content that the person actually produced, whereas generative systems create new statements and reactions in the person's name.

Against this background, the paper conceptualizes perceived authenticity as the perceived fit between an AI-generated representation and the remembered person. It develops a multidimensional framework for specifying the aspects through which this fit is established and translates these dimensions into an AI Afterlife agent prototype.

5. Dimensions of Authenticity

Based on the literature reviewed, perceived authenticity is conceptualized as the perceived fit between an AI-generated representation and the remembered person. To operationalize this construct, the reference person is represented along seven dimensions: *Traits*, *Facts*, *Behavior*, *Decision Logic*, *Values*, *Boundaries*, and *Style*. The dimensions capture complementary aspects through which close others may recognize a specific person (see Table 1).

Table 1: Dimensions of Authenticity

Dimension	Definition and theoretical grounding	Prototype operationalization
Traits	Relatively stable personality characteristics and dispositional tendencies; grounded in personality research (Costa & McCrae, 1992).	Baseline self-description: *"Describe briefly what kind of person you generally were in everyday life"*
Facts	Verifiable biographical information anchoring the representation in the person's life history; informed by digital-legacy and AI-afterlife research (Doyle & Brubaker, 2023; Lei et al., 2025).	Baseline biographical self-description.
Behavior	Recurring observable patterns of action and reaction in situations such as stress, conflict, routine, or emotional interaction (Kernis & Goldman, 2006; Voinea et al., 2026).	Baseline response to difficulty/tension: *"How did you normally react when things became difficult or tense?"*
Decision Logic	Recognizable reasoning and decision rules underlying action, such as handling uncertainty, intervention vs. restraint, and prioritization (Kernis & Goldman, 2006; Voinea et al., 2026).	Baseline decision scenario: *"How did you normally proceed when you were uncertain or had to make a decision?"*
Values	Enduring normative priorities and evaluative commitments that guide judgments and trade-offs (Schwartz, 1992).	Baseline value question: *"What was particularly important to you in life?"*
Boundaries	Personal limits concerning privacy, disclosure, intimacy, tone, appropriate interaction, and topics/actions the person would reject; grounded in Communication Privacy Management (Petronio, 2002).	Baseline boundary question: *"What clearly would not have fitted in the way someone interacted with you?"*
Style	Recognizable linguistic and communicative expression, including wording, tone, directness, humor, formality, and typical forms of address (Linden, 2025).	Baseline reunion interaction: *"How would you normally speak with me if we had not been in contact for a while?"*

6. Conception of the AI Afterlife Agent

The conception of the research artifact follows two central requirements. First, the (living) reference person must be captured in a structured way along the seven authenticity dimensions. Second, the subsequent interaction must allow individual dimensions to remain stable or be varied in a controlled manner. These requirements lead to a two-stage architecture consisting of an assessment phase and a simulation phase, as illustrated in Figure 1. The technical implementation of the research artifact was carried out using the PROMISE framework (Wu et al., 2026/2023).

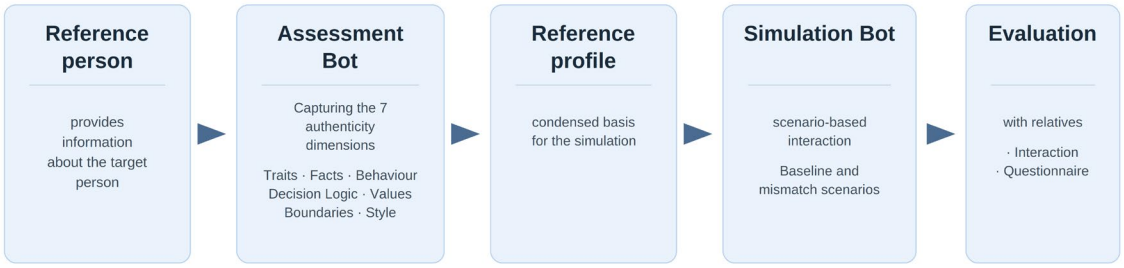

Figure 1: Conception of the AI Afterlife agent prototype

6.1 Assessment Phase

The Assessment Bot collects structured information about the living reference person along the seven authenticity dimensions. Each dimension is represented by a separate interaction state in the PROMISE framework. The collected information is condensed into a reference profile that serves as the binding basis for the subsequent simulation.

6.2 Simulation Phase

The Simulation Bot uses the structured reference profile as the binding source for person-specific responses across standardized scenario states (for details, see Steiner & Moser, 2026). The complete assessment dialogue was not used as simulation input; instead, only the condensed reference profile and the current scenario prompt informed response generation. The bot was instructed not to extend or reinterpret the profile. Seven baseline scenarios generated profile-consistent responses, while two mismatch scenarios deliberately altered either *Facts* or *Style*. In mismatch states, only the target dimension was changed while the remaining dimensions were kept as stable as possible. Responses were restricted to one short, natural answer to improve comparability and reduce unintended cues.

7. Evaluation

7.1 Study Design and Sample Description

The evaluation of the research artifact was conducted as a scenario-based within-subjects study with repeated measures. The final evaluation run comprised nine scenarios: seven baseline scenarios and two mismatch scenarios. The mismatch scenarios deliberately violated selected authenticity dimensions, namely *Style* and *Facts*, by introducing responses that should not fit the reference person.

The study included seven family contexts. Each context consisted of one reference person, who had previously created the profile, and two closely related participants who evaluated the simulation. This resulted in 14 completed questionnaires. Since each participant assessed all nine scenarios of the same simulated reference person, a total of 126 scenario-based evaluations were included in the analysis. However, the evaluations constitute of repeated observations and are not independent cases.

Participants interacted individually with the simulation bot and evaluated each response immediately after the respective scenario using a standardized questionnaire (for details, see Steiner & Moser, 2026). After each scenario, participants completed Likert-scale items and open-ended questions. After completing all scenarios, they additionally answered prioritization tasks concerning the relative importance of the authenticity dimensions. The questionnaire used seven-point Likert items ranging from 1 = strongly disagree to 7 = strongly agree. After each scenario, participants evaluated five items. The first four items were identical across all scenarios and captured general perceived authenticity:

1. This response sounds like the person.
2. Word choice and tone fit the person.
3. Content and emphasis fit the person.
4. Something about this response seems untypical or inconsistent.

The fourth item was negatively oriented and was therefore interpreted in reverse when calculating the authenticity score. The fifth item was scenario-specific and captured the targeted authenticity dimension of the respective scenario.

The authenticity index was calculated as a scenario-level aggregate measure of perceived authenticity. It combines the four general authenticity items with the respective dimension-specific item. Before aggregation, the negatively oriented fourth item was recoded so that higher values consistently indicate higher perceived authenticity across all items. The index therefore represents the mean perceived authenticity of a given response on a scale from 1 to 7, with higher values indicating stronger perceived authenticity. As a descriptive reliability check, Cronbach's alpha ranged from .787 to .918 across the baseline scenarios and was .958 for the Style mismatch and .951 for the Facts mismatch. Given the small sample, these values are interpreted only as indicators of internal consistency within the present dataset.

Only participants with a close relationship to the reference person and sufficient memory confidence were included (see Table 2), as the study depended on their ability to judge whether the simulation corresponded to their lived experience of that person. Participation was voluntary, and participants could stop the interaction with the simulation bot at any time. All participants were recruited through personal family and relationship contexts.

Table 2: Study participants

Relationship to the reference person	n	%
Other relatives	4	28.6
Siblings	4	28.6
Parent	2	14.3
Partner	2	14.3
Close friend	2	14.3

7.2 Results

7.2.1 Comparison of the authenticity index across scenarios

The descriptive results show a clear pattern (for details, see Steiner & Moser, 2026). Table 3 summarizes the results across all scenarios. Baseline authenticity was consistently high (M = 6.04-6.66), whereas both mismatch scenarios received lower ratings. The *Facts* mismatch showed the largest descriptive reduction (M = 4.17, SD = 2.33), while the *Style* mismatch remained closer to its corresponding baseline (M = 5.67, SD = 0.81). Given the small exploratory sample, these comparisons are interpreted descriptively.

The open-ended responses support this pattern. In the Facts-Mismatch scenario, participants identified concrete biographical inconsistencies. Examples include: "X ist nicht in Bern aufgewachsen, wie es in der Antwort heisst" ["X did not grow up in Bern, as stated in the answer"] (F6_A2), "X hat nie als Pflegefachfrau gearbeitet, so wie es in der Antwort heisst" ["X never worked as a nurse, as stated in the answer"] (F4_A2), and "Die Antwort, dass X Polizist ist, stimmt nicht" ["The answer saying that X is a police officer is incorrect"] (F7_A2). These comments show that incorrect biographical information was experienced as a direct authenticity violation. In the Style-Mismatch scenario, participants focused on wording, tone, and linguistic fit. One participant stated: "Den Begriff 'schnurze' hätte X nie benutzt" ["X would never have used the term 'schnurze'"] (F5_A2). Another described the response as "ein wenig zu kalt und generisch" ["a little too cold and generic"] (F5_A2) and added that it did not sound like the person. These responses indicate that stylistic authenticity depends on recognizable language use, emotional register, and tone.

Table 3: Authenticity Index by Dimension

Dimension	Condition	M	SD
Traits	Baseline	6.44	0.47
Facts	Baseline	6.49	0.40
Behaviour	Baseline	6.64	0.39
Decision Logic	Baseline	6.07	0.62
Values	Baseline	6.66	0.38
Boundaries	Baseline	6.60	0.46
Style	Baseline	6.04	0.92
Style	Mismatch	5.67	0.81
Facts	Mismatch	4.17	2.33

7.2.2 Prioritization of authenticity dimensions

Additional insight into the relative importance of the authenticity dimensions comes from the final forced-choice question of the questionnaire. In this task, participants were asked to select the two aspects that were most important for their overall authenticity judgement. The prioritization task was fully analyzable across all responses (n=14x2). The results are visualized in Table 3.

Table 3: Prioritization ratings of Authenticity Dimensions

Dimension	n	%
Traits	7	50.0
Facts	11	78.6
Behavior	0	0.0
Decision Logic	0	0.0
Values	4	28.6
Boundaries	0	0.0
Style	6	42.9

The results show a clear hierarchy. *Facts* were selected most frequently, with 11 mentions corresponding to 78.6% of participants. This indicates that biographical correctness and factual consistency played the strongest role in participants' overall evaluation of authenticity. *Traits* were the second most frequently selected dimension, with 7 mentions or 50.0%. This suggests that stable personality characteristics were also highly relevant for judging whether the simulated response matched the reference person. *Style* was selected by 6 participants, corresponding to 42.9%. This shows that linguistic expression, tone, and recognizable wording were important for a substantial share of participants. *Values* were selected by 4 participants, corresponding to 28.6%. This indicates that the perceived alignment with the reference person's moral orientation and core beliefs also contributed to the overall authenticity judgement, although less frequently than *Facts*, *Traits*, and *Style*.

In contrast, Behavior, *Boundaries*, and *Decision Logic* were not selected in the forced-choice task. This does not necessarily mean that these dimensions were irrelevant for perceived authenticity. Their high ratings in the baseline scenarios suggest that they were perceived as fitting in the evaluated responses. However, when participants had to prioritize only two aspects, they placed stronger emphasis on factual correctness, personality traits, linguistic style, and values.

8. Discussion

The evaluation indicates that the prototype was generally successful in generating responses that participants perceived as authentic. The consistently high Authenticity Index values across the seven baseline scenarios suggest that the structured reference profile provided a suitable basis for simulating the reference person in a recognizable and coherent way.

The results also show that the targeted manipulations worked as intended. Both mismatch scenarios received lower ratings than the baseline scenarios and prompted explicit comments from participants. This indicates that the introduced deviations were noticeable and affected perceived authenticity. The strongest effect was found for *Facts*. The Facts-Mismatch scenario received the lowest Authenticity Index, and *Facts* were also selected most frequently in the prioritization task. Factual correctness therefore appears to be a central condition for perceived authenticity, as biographical errors directly disrupt the perceived connection between the agent and the remembered person. *Style* was also relevant, although its effect was less pronounced. The Style-Mismatch scenario was rated lower than the baseline scenarios, and *Style* was selected by several participants as important for their overall judgement. This suggests that tone, wording, and linguistic recognizability are salient cues in interaction. *Facts* and *Style* may be particularly influential because they are immediately visible: participants can quickly notice whether information is correct and whether the response sounds like the reference person. The strong relevance of *Facts* also suggests that factual correctness remains a basic condition of perceived authenticity, even where the literature emphasizes broader relational, behavioral, or value-based aspects. It may function as a hygiene factor: when factual information is correct, it may receive little explicit attention, but when it is wrong, the mismatch becomes highly salient and can undermine the authenticity of the entire representation.

By contrast, *Decision Logic*, *Boundaries*, and *Behavior* were not selected in the prioritization task. This does not imply that they are irrelevant, as their baseline ratings were high. However, these dimensions may be less directly visible in short scenario-based interactions. They may also overlap with broader categories. *Behavior* may be interpreted as part of *Traits*, while *Decision Logic* and *Boundaries* may be understood through *Values*. This suggests that participants may not distinguish all seven dimensions as clearly as the theoretical framework does.

Overall, the findings provide preliminary support for a dimensional conceptualization, while the seven theoretically derived dimensions require further empirical differentiation. The seven-dimensional model was useful for designing and controlling the prototype, but users may rely on fewer, more immediately recognizable cues when judging authenticity.

9. Conclusion

This paper examined how perceived authenticity in AI Afterlife agents can be conceptualized, implemented, and evaluated. Based on the literature, seven authenticity dimensions were derived and translated into a scenario-based AI Afterlife agent prototype.

The evaluation indicates that the prototype generally worked as intended. The high Authenticity Index values in the baseline scenarios show that a structured reference profile can support the simulation of a specific person in a recognizable and coherent way. The chosen approach, in which the reference person voluntarily creates their own profile during their lifetime, is also useful from an ethical perspective because it strengthens consent and control.

The research questions can largely be answered. First, relevant authenticity dimensions could be theoretically derived from existing literature. Second, all seven dimensions were operationalized in the prototype, while Facts and Style received initial experimental evidence through targeted mismatch manipulations. However, the findings suggest that the seven-dimensional model may need refinement, as some dimensions may overlap or may not be clearly distinguishable for participants. Third, the results show that *Facts* have particular relevance for perceived authenticity. Factual correctness may function as a basic hygiene factor: it attracts limited attention when present but becomes highly salient when violated. Accordingly, factual inaccuracies were clearly recognized and had the strongest negative effect on authenticity.

The evaluation should be interpreted as an exploratory prototype study, since it is limited by its small sample and by the fact that only two dimensions, *Facts* and *Style*, were deliberately manipulated. Future research should therefore examine the results in more diverse samples and explore whether deviations in other dimensions would be recognized in a similar way. In addition, longitudinal studies are needed to understand how perceived authenticity, emotional effects, and potential risks develop over repeated interaction with an AI Afterlife agent.

Ethics Declaration: Participation was voluntary and based on informed consent; participants could skip questions, take breaks, or discontinue the study at any time, and the interaction was paused or terminated in case of apparent distress. Participants were not recruited during an acute bereavement situation. All reference persons were living and provided their profile information voluntarily; evaluators therefore interacted with

simulations of living close others rather than deceased persons. The study presented the system explicitly as an AI-based simulation and not as the represented person.

AI Declaration: Generative AI tools were used to support language editing, structural refinement, and the development of the research prototype. All AI-assisted content was reviewed and revised by the authors, who remain fully responsible for the study design, analysis, interpretation, and final manuscript.

References

Battisti, D. (2025). Second-Person Authenticity and the Mediating Role of AI: A Moral Challenge for Human-to-Human Relationships? *Philosophy & Technology*, *38*(1), 28. https://doi.org/10.1007/s13347-025-00857-w

Costa, P. T., & McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of Personality Disorders*, *6*(4), 343–359. https://doi.org/10.1521/pedi.1992.6.4.343

Doyle, D. T., & Brubaker, J. R. (2023). Digital Legacy: A Systematic Literature Review. *Proceedings of the ACM on Human-Computer Interaction*, *7*(CSCW2), 1–26. https://doi.org/10.1145/3610059

Figueroa-Torres, M. (2024). Affection as a service: Ghostbots and the changing nature of mourning. *Computer Law & Security Review*, *52*, 105943. https://doi.org/10.1016/j.clsr.2024.105943

Fu, K., Ye, C., Wang, Z., Liu, Z., Wu, M., & Yuan, Y. (2025). Ethical dilemmas and the reconstruction of subjectivity in digital mourning in the age of AI: An empirical study on the acceptance intentions of bereaved family members of cancer patients. *Frontiers in Digital Health*, *7*, 1618169.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–105. https://doi.org/10.2307/25148625

Hollanek, T., & Nowaczyk-Basińska, K. (2024). Griefbots, Deadbots, Postmortem Avatars: On Responsible Applications of Generative AI in the Digital Afterlife Industry. *Philosophy & Technology*, *37*(2), 63. https://doi.org/10.1007/s13347-024-00744-w

Jiménez-Alonso, B., & Brescó de Luna, I. (2023). Griefbots. A New Way of Communicating With The Dead? *Integrative Psychological and Behavioral Science*, *57*(2), 466–481. https://doi.org/10.1007/s12124-022-09679-3

Jongman-Sereno, K. P., & Leary, M. R. (2019). The Enigma of Being Yourself: A Critical Examination of the Concept of Authenticity. *Review of General Psychology*, *23*(1), 133–142. https://doi.org/10.1037/gpr0000157

Kernis, M. H., & Goldman, B. M. (2006). A Multicomponent Conceptualization of Authenticity: Theory and Research. In *Advances in Experimental Social Psychology* (Vol. 38, pp. 283–357). Academic Press. https://doi.org/10.1016/S0065-2601(06)38006-9

Kidd, J., & McAvoy, E. N. (2025). Synthetic afterlives: Deathbots as affective infrastructures of memory. *Memory, Mind & Media*, *4*, e16. https://doi.org/10.1017/mem.2025.10013

Lei, Y., Ma, S., Sun, Y., & Ma, X. (2025). "AI Afterlife" as Digital Legacy: Perceptions, Expectations, and Concerns. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–18. https://doi.org/10.1145/3706598.3713933

Li, J., Galley, M., Brockett, C., Spithourakis, G., Gao, J., & Dolan, B. (2016). A Persona-Based Neural Conversation Model. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 994–1003. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). https://doi.org/10.18653/v1/P16-1094

Li, X., & Sung, Y. (2021). Anthropomorphism brings us closer: The mediating role of psychological distance in User–AI assistant interactions. *Computers in Human Behavior*, *118*, 106680. https://doi.org/10.1016/j.chb.2021.106680

Linden, K. (2025). The Talking Dead: Voice-Cloning Yourself to AI Afterlife Services. : Analysis of Representation Authenticity and Social Implications. *Thanatos*, *13*(2). https://doi.org/10.57124/thanatos.160215

Liu, Y., Wei, W., Liu, J., Mao, X., Fang, R., & Chen, D. (2022). Improving Personality Consistency in Conversation by Persona Extending. *Proceedings of the 31st ACM International Conference on Information & Knowledge Management, CIKM '22*, 1350–1359. https://doi.org/10.1145/3511808.3557359

Lu, J., Yang, M., Cui, Y., Liu, Y., Wong, T. K. S. K. S., & Hu, Y. (2026). Exploring public perceptions of artificial intelligence in bereavement support: A qualitative study of griefbots. *Frontiers in Digital Health*, *8*. https://doi.org/10.3389/fdgth.2026.1852514

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *The Academy of Management Review*, *20*(3), 709–734. https://doi.org/10.2307/258792

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*, *19*(2), 98–100. https://doi.org/10.1109/MRA.2012.2192811

Nunes, J. C., Ordanini, A., & Giambastiani, G. (2021). The Concept of Authenticity: What It Means to Consumers. *Journal of Marketing*, *85*(4), 1–20. https://doi.org/10.1177/0022242921997081

Öhman, C., & Floridi, L. (2018). An ethical framework for the digital afterlife industry. *Nature Human Behaviour*, *2*(5), 318–320. https://doi.org/10.1038/s41562-018-0335-2

Pandey, P., & Rai, A. K. (2024). Analytical modeling of perceived authenticity in AI assistants: Application of PLS-predict algorithm and importance-performance map analysis. *South Asian Journal of Business Studies*, *14*(1), 1–30. https://doi.org/10.1108/SAJBS-04-2023-0102

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Petronio, S. (2002). *Boundaries of Privacy: Dialectics of Disclosure*. State University of New York Press. https://www.jstor.org/stable/jj.18255126

Riccardo, D., Keller, T., & Moser, M. (2026). *Directed Recognizability in AI-Biographical Conversational Agents: Design and Evaluation in the Digital Afterlife Context* (4b8gy_v2). PsyArXiv. https://osf.io/preprints/psyarxiv/4b8gy_v2/

Schwartz, S. H. (1992). Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 25, pp. 1–65). Academic Press. https://doi.org/10.1016/S0065-2601(08)60281-6

Sonlu, S., Güdükbay, U., & Durupinar, F. (2021). A Conversational Agent Framework with Multi-modal Personality Expression. *ACM Transactions on Graphics (TOG)*, *40*(1), 7:1-7:16. https://doi.org/10.1145/3439795

Sousa, T., Bizarrias, F., Raposo, L., & Cruz, D. (2025). When AI becomes afterlife: Psychological factors driving acceptance of thanabots services. *Journal of Services Marketing*, *40*. https://doi.org/10.1108/JSM-01-2025-0052

Spitale, G., & Germani, F. (2025). The making of digital ghosts: Designing ethical AI afterlives. *Ethics and Information Technology*, *28*. https://doi.org/10.1007/s10676-026-09910-4

Steiner, J., & Moser, M. (2026). *Authenticity Dimensions in AI Afterlife agents*. https://osf.io/am3g7/overview

Voinea, C., Mann, S. P., Savulescu, J., & Earp, B. D. (2026). Digital Doppelgängers, Human Relationships, and Practical Identity. *Bioethics*, *40*(5), 435–444. https://doi.org/10.1111/bioe.70026

Wu, W., Heierli, J., Gavagnin, E., & de Spindler, A. (2023). *PROMISE* [HTML]. https://github.com/zhaw-iwi/promise (Original work published 2026)

Zhang, Z., Dou, D., & Liu, X. (2025). Co-construction with deathbots: A form of existentialist practice. *Humanities and Social Sciences Communications*, *12*(1), 537. https://doi.org/10.1057/s41599-025-04876-y