My Fate Is to Die Young, But to Live Forever in Song: Echeloned Design Science Research to a Digital-Me Expert System Design

Petri Parvinen
University of Helsinki
petri.parvinen@helsinki.fi

Osmo Mattila
Aalto University
osmo.mattila@aalto.fi

Mohammad Soltanieh-ha
Boston University
msoltani@bu.edu

Tuure Tuunanen
University of Jyväskylä
tuure@tuunanen.fi

Abstract

This study explores the developmental trajectory and implementation of an artificial intelligence (AI) -based expert system. Following and extending the echeloned design science research approach, the process examines how the nature of system development changes when the team designing the system shifts from using AI to produce expert artifacts to using AI as the core technology to continuously run digital clones of the expert. We find a radically transformed nature of design and development work and numerous modes of evaluation echelons. Methodologically, the results contribute to DSR in information systems by articulating AI's role not only as a tool for artifact creation, but as a co-evolving actor within the design process, and the changes that entails to the DSR approach. Theoretically and empirically, we discover and develop the notion of digital-me, hypothesizing the benefits of conceptualizing AI as extending real humans, not building humanlike artificial agents, as the prime use of AI, particularly in user acceptance of expert systems.

Keywords: Artificial intelligence, digital-me, video avatar, text clone, disclaimer, expert system, echeloned design science research.

1. Introduction

Artificial intelligence (AI) agents (e.g., Golovianko et al., 2023; Wang et al., 2024) have been central to AI research because they can automate tasks that used to require human expertise. In this paper, we look at expertise by asking: what if, instead of being replaced, an expert created a digital version of themselves to make their knowledge and persona more widely available? We call this approach digital-me. We define digital-me as with video, voice and substantive content that replicates a human expert interacting to give advice in a given domain.

AI opens opportunities to utilize knowledge and appearance-related patterns in novel ways. Currently, chatbots are probably the most common way to utilize large language models (LLMs) in customer interface. For many users, however, a unified interface that feels more like interacting with a consistent digital persona might be more compelling. Chen et al. (2024) argue that human-personalization in LLMs is still a largely unexplored area. Further, a limited model may be more cost-effective to use than a general model. This is a crucial consideration, as the operational costs of AI models have tended to rise (Annepaka & Pakray, 2025), in tandem with their complexity and ability to perform human-like expert tasks. According to Agiollo et al. (2024), by decomposing complex tasks into subtasks and assigning each to the most cost-effective AI agent, expert systems have been seen to achieve significant efficiency gains.

Beyond knowledge, we argue that expertise is inherently linked to the individual. Accordingly, we examine the cloning of appearance-related features, including an expert's voice and looks. This paper documents the creation of an AI-driven digital extension of a human expert, enabling the human to scale according to client needs. By using the echeloned design science research (^{e}DSR) approach by Tuunanen et al. (2024), the development processes of four AI-enabled expert system features are documented. Two development paths are followed: digitalizing appearance by cloning voice and video presence, and digitizing knowledge by making a text clone.

2. Background

The literature explaining processes of conducting artifact- and theory-driven design science research (DSR) is attempting to keep pace with evolving software and system development practices. One of the latest updates has been the abandonment of the traditional version-based approach of the DSR approach (Hevner et al., 2004; Peffers et al., 2007) and complementing it with ^{e}DSR (Tuunanen et al., 2024), which breaks the fashion of conducting DSR, more towards, e.g., agile development, where pivoting is essential and a development direction may change at a moment's notice.

With the large-scale introduction of an LLM as the background of interactive agents, attention has turned to designing such constantly evolving, learning, and adapting agents to solve expert subtasks (e.g., Wang et

URI: https://hdl.handle.net/10125/111751
978-0-9981331-9-5
(CC BY-NC-ND 4.0)

al., 2024; Agiollo et al., 2024). According to Liang and Wei (2024), a multi-agent ecosystem, resembling current application markets, is expected to emerge to meet complex expertise needs.

Digital extensions of real humans can be a part of a network of expert systems. They may increase credibility but also limit the user's expectations to a number of use cases, making it easier to manage and maintain the models. It is becoming increasingly difficult to make a difference between real and cloned voices (Mcuba et al., 2023), faces (Miller et al., 2023) and texts (Fiedler & Döpke, 2025). This paper addresses this research gap by documenting a multi-pronged research, development, and implementation process for a digital-me agent of a real expert human.

At the time the study was initiated, no ready-made digital-me agents were available, merging both digital presence and knowledge models. In this study, ^{e}DSR guides a process involving fast-paced advances in technology, human-computer interaction, and user expectations in the exploration of digitally cloning a human expert presence. Avatars have generally been evaluated based on their human-like features. Depending on the features that were tested, we utilized the same measure that in earlier research have been used to measure such features related to avatars, (Seymor et al., 2021; Yuan & Dennis, 2019), AI assistants (Hu et al., 2021), and humans experienced via video or other remote communication (Rubin et al., 1985; Judd et al., 2005; Epley et al., 2007; Hartmann & Goldhoorn, 2011). Based on previous research, realism in features tends to be positively perceived (e.g., Ferst et al., 2021), but mismatches, such as an unreal voice displayed with a realistic virtual human, can be unappealing (Higgins et al., 2022). The hypothesis is that by emulating aspects of human behavior and communication, AI systems can become more intuitive and engaging for users.

Currently, there is no off-the-shelf technology available to integrate a functioning clone of an expert human to be represented in live interaction, featuring video-type avatar person, with a natural voice and using the data and documentation of the expert's knowledge. Thus, in this research, we employ a multi-pronged approach to investigate user acceptance of these features. Specifically, our research questions are:

1. How does the emulation of human behavior and communication by AI drive user acceptance of the digital-me?
2. How does the emulation of knowledge related to an individual expert drive the use of the digital-me?
3. How does the nature of system development change when constantly updating LLM models that represent the core of the expert system's operation?

3. Research Design

In this study, we conducted a multi-year DSR project following the ^{e}DSR methodology (Tuunanen et al., 2024). This approach enabled us to iteratively explore and refine the features through multiple design echelons, each addressing the evolving technological and user experience challenges of digital cloning an expert human (see, e.g., Golovianko et al., 2023). It distinguishes five types of design echelons: (1) problem analysis, (2) objectives and requirements definition, (3) design and development, (4) demonstration, and (5) evaluation. Our project iteratively cycled through these design echelons to adapt to rapid changes in AI capabilities and user expectations. Each echelon used distinct validation criteria (VC), which is outlined with each echelon.

The main differences between DSR and ^{e}DSR are that firstly, it enables empirical observations of micro- and macro-level echelons without burdening the process with extensive evaluation phases in cases where they are not needed. Secondly, it allows for greater flexibility in switching between features. This helps orchestrate practical development without getting stuck on developing extensively challenging features or an occasional lack of development resources.

This study is, in large part, explorative. Consequently, not all development cycles culminate in a typical evaluation phase. Instead, system assessments involving domain experts or end-users were often sufficient to facilitate further development. These early-stage evaluations suggested that postponing further work until more accessible and implementable technological solutions become available would be a more strategic course of action.

The voice clone was the first feature developed in 2023. The second feature, video clone, was created in 2024 and integrating it with the previously created voice clone. The third feature, video dubbing, was conducted by testing a video translation feature. The fourth feature, text clone, targeted making the character interactive by improving a generic LLM, using the retrieval augmented generation approach by feeding the model with material previously published by the person in question. Each feature developed in this study was demonstrated and evaluated. The pace of iterations accelerated along with interactivity when moving to chatbot.

4. Designing and Evaluating the Digital-Me Expert System

The project overview is presented in Figure 1. The voice clone was covered in design echelons (E) E1-E5. Once ready, it was integrated into both the video cloning and video dubbing features. The development of video clones is covered in echelons E6-E8, and video dubbing in E9-E11. This step proceeded to the user evaluation phase, where the humanity- and trust-related features of AI-modified and AI-generated videos were compared. Text clone development is covered in echelons E12 to E17. and it was developed by ten rapid cycles that were documented in a detailed activity log with timestamps, persons involved, and detailed content of the activity (full text or transcription whenever available) and theoretical observations/interpretations.

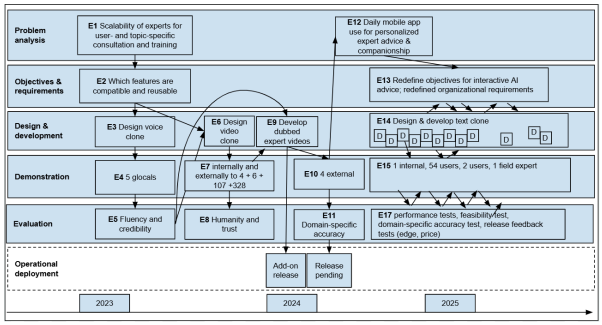


Figure 1. The overview of the project. AI-created assets used 2023-2024, LLM-driven operation commenced in January 2025.

Because internal evaluations became such an integral part of creating the text clone, we added an Internal Evaluation phase before the Demonstration phase. Furthermore, since some features were deployed directly to production, we also included an Operational Deployment phase at the end of the process.

Problem Analysis. One topical approach to incorporating AI systems into practical business use cases is the use of task-specific, specialized AI agents. This approach emphasizes efficient resource allocation: each task is delegated to the most cost-effective agent capable of performing it. However, while this model optimizes computational and operational efficiency, it often overlooks the user experience dimension.

We explore the potential of human-like communication as a means to enhance the intelligibility and accessibility of AI functionalities. Although this is only one possible path for AI user interfaces, it led us investigate which elements of human identity and interaction can be meaningfully digitized and operationalized within AI systems.

Two parallel development tracks were pursued: Our development process started from cloning the "interface" or “humanity” of a character by focusing on the external, perceptual aspects of interaction. Later, another development path was started. It was related to the "content" of a character, emphasizing internal knowledge while being aware that personality would be lacking when using only knowledge-intensive training data instead of e.g., personal messages. However, these efforts aim to determine which components are essential for effective interaction and which are technically feasible to implement. A critical insight from this work is that, while many features are theoretically feasible and demonstrable in controlled environments, their practical deployment often faces significant challenges. By applying the [e]DSR method, these challenges were structured into the following problem-analysis-related echelons:

E1 (VC: scalability): Scalability of Experts – There is a growing need for scalable AI systems that can provide expert-level consultation and training tailored to specific users and topics. Since we followed a path of emulating an expert, problem analyses focused on digitizing the expert's presence. Being fully aware of the forthcoming challenges of creating a character that feels both interactive and natural, the first use cases aimed to publish various pre-recorded materials, such as video courses given by experts.

E12 (VC: usability): Daily Mobile App Use – Based on the internal evaluation results of E10, the research team concluded that integrating the expert AI agent into daily mobile app use opened an entirely new problematization for the project. App use was perceived as necessary for the value creation avenue, as it delivered personalized expert advice and daily companionship. However, this was seen to require not only technical robustness but also an altogether new understanding of user preferences, emotional cues, and contextual relevance. Even though the video clones moved into deployment, the project also pivoted to create a text clone based on a private knowledge repository made available by the expert.

In conclusion, while the modular, task-specific architecture of AI agents offers clear advantages in efficiency, the human-centric design of these systems remains a critical frontier to address. Bridging this gap requires interdisciplinary efforts that combine technical innovation with insights from human-computer interaction, cognitive science, and design.

Objectives and requirements. Following the identification of E1: Scalability of Experts, a research question emerged: E2 (VC: feature scalability): Which

features are compatible and reusable? This question addresses the modularity and interoperability of AI-generated components across different use cases and platforms. In the context of developing scalable expert personas, it is essential to determine which features—such as voice, visual identity, and behavioral patterns can be efficiently reused or adapted across multiple agents or applications. This line of inquiry is particularly relevant for the development of voice clones and video clones, which require a careful balance between fidelity, flexibility, and integration potential. To address these challenges, a multilingual approach was adopted, since the individual whose voice was cloned is a native speaker of a language underrepresented in AI training data. Ensuring that the clone could operate effectively beyond English was a key requirement.

From a practical standpoint, two validation criteria guide the selection and development of such features: (a) Commercial viability: Features should be designed for rapid deployment and monetization, either as standalone services (e.g., publishing audio books of the books the expert has published earlier) or as integrated components in broader applications (e.g., virtual assistants, training simulators). (b) Interoperability: Features must be compatible with a wide range of models and systems, enabling integration with existing other features under development.

The exploration of reusable features thus serves both strategic and technical purposes: it accelerates time-to-market while laying the groundwork for scalable, cross-domain AI agents that can maintain coherent identity and functionality across diverse contexts.

In the text-clone approach, by building on E12 and the next step is E13 (VC: efficacy): Redefine objectives for interactive AI advice, that as a result of development iterations in E14 (VC: efficacy), also included organizational requirements. As AI agents become persistent, integrations to the knowledge base of an expert become relevant.

The reason for modifying a generic LLM was to enhance its ability to produce personified responses. By combining the model's learned language patterns with retrieval from earlier publications by the expert, the system was intended to provide answers that would be more grounded in research-based content within a limited area, but beyond the original training data of the generic LLM.

Design and Development. From the analysis in E2: Which features are compatible and reusable, two concrete design challenges emerge: E3 (VC: quality): Design voice clone and E6 (VC: quality): Design video clone. These represent critical modalities for creating reusable, identity-preserving features for the digital expert. Voice and video are not only high-impact channels for user engagement but also technically complex, requiring careful consideration of humanness-related features. We used these as the validation criteria. Designing a voice clone involves selecting features such as tone, prosody, and emotional range that can be consistently reproduced across contexts while remaining lightweight enough for real-time use. Similarly, video cloning must strike a balance between visual realism and modularity, enabling expressive, context-aware avatars that can be seamlessly embedded into diverse platforms. Both modalities must be designed with reusability in mind, ensuring that cloned assets can be adapted to new domains, updated over time, and combined with other AI components without loss of coherence or performance.

After testing open-source tools, we chose ElevenLabs for its high-quality voice cloning and easy API integration. While the clone worked well, producing natural voice messages often required multiple iterations to refine pronunciation and emphasis, raising concerns about usability in interactive scenarios needing expressive versatility.

Similarly, video cloning should enable expressive, context-aware avatars that can be embedded into diverse platforms. Both modalities must be designed with reusability in mind, ensuring that cloned assets can be adapted to new domains, updated over time, and combined with other AI components without loss of coherence or performance. A video clone was created by using Synthesia’s studio avatar feature. By using this avatar, we were able to start producing video lectures based on text prompt,s but still having the expert to write the manuscripts.

E9 (VC: utility): Develop dubbed expert videos that emerged from a need to evaluate the performance of video clones. Since we recognized that AI-dubbed video translation falls between a conventional and an AI-generated version, we decided to include it as a distinct feature in the process.

E14 (VC: quality): Design and development of the text clone was derived from E13. We used a generic ChatGPT model as a baseline model and it was augmented by publications (16 journal articles, six books, two consultative texts) and training material (38 powerpoint decks with 290 powerpoints) authored by the expert by that time and its operation was prompted by a dedicated 1500 word prompt on how to use and prioritize private knowledge repository, how to amend it with generic LLM knowledge when necessary and how to address giving expert advice stylistically.

Demonstration. Demonstrations varied from rapid internal tests to comparative experiments. In E4 (VC: acceptance), the voice clone was demonstrated to five people speaking Filipino, Finnish, Swedish, and English. They were either sent a voice message or asked in person to listen to an audio message generated by using the voice clone.

In E7 (VC: acceptance), a video clone was first demonstrated internally among four researchers. To make comparison more thoughtful, a video translated by AI was created, as well as a conventional video. After small improvements, these three videos were demonstrated to five students and one external researcher. Thereafter, the study was expanded to cover 107 students and later on 328 testers recruited through MTurk.

AI-dubbed expert videos were found to be useful as such. In one use case, they went directly from development in E9 to production. In another use case in E10 (VC: realism), they were evaluated by four language specialists. In this case, they commanded Estonian, Arabic, Finnish, and English.

With the text clone, however, more demonstrations were needed in E15 (VC: usability). There were two main reasons for that. Firstly, the review rounds required extensive involvement from the expert whose input was being replicated. Secondly, text clone development in this case was a much more iterative process compared to audio and video cloning that were heavily dependent on existing services. Text clone was first demonstrated to five team members, a self-test by the expert, and as an internal performance test that was conducted by two people participating in the development of the bot. The four single users were requested to use the system in their daily operations, requiring the expert's advice through the dedicated GPT text clone. It was then tested externally as a sales demo for business-to-business use in, and as a demo for a field expert. Based on these tests, it was demonstrated to two sales experts. Finally, it was demonstrated to a professional in education technology.

Evaluation. Evaluations focused on two main categories. The digitalized appearance was assessed based on validation criteria based on human-centricity scales, while the knowledge clone was evaluated on usability criteria.

In E5 (VC: efficacy), the voice clone was evaluated by five people who knew the cloned person beforehand. They were interviewed about their opinions on the credibility and pronunciation of the voice clone. Pronunciation was described as slightly monotonic, but overall, the voice clones were described to closely resemble the specific individual. Evaluations in E8 (VC: efficacy) involved a comparative analysis of three types of video content: an AI-generated video, an AI-dubbed video, and a conventional, unaltered video. In addition to comparing the content types themselves, the study also examined how user experience varied depending on whether participants were informed about the use of AI in the content creation process. The goal was to assess not only perceptual and qualitative differences between the video types but also to understand the impact of transparency regarding AI involvement. The aim of the user evaluations was to test the effect of A disclaimer on the experience and to compare which videos performed the best in general.

We used survey studies and previously established scales for this purpose: humanness (Epley et al., 2007; Jia et al., 2017; Yuan & Dennis, 2019), affinity (Seymor et al., 2021), warmth (Hu et al., 2021; Judd et al., 2005; Zhou et al., 2019), competence Hu et al., 2021; Judd et al., 2005; Zhou et al., 2019; Aaker, 1997) trustworthiness (Ohanian, 1990), parasocial interaction (Rubin et al., 1985; Hartmann & Goldhoorn, 2011) and brand trust (Patrizi, Seric, & Vernuccio, 2023). Intention to act and attitude towards AI were measured using items from Komiak and Benbasat (2006) and Stein et al. (2024). Given the straightforward goal of comparing treatment groups and the relatively small sample sizes per group, composite variables were formed using the sum variable technique based on the literature. All scales, except for Attitude towards AI, performed as expected for students (n = 107), with Cronbach's alpha values ranging from 0.82 to 0.93.

These same scales were employed in the second evaluation phase with MTurk participants (n = 328), and reliability remained comparable, with alpha values ranging from 0.84 to 0.94. In the first evaluation, data was not normally distributed. Small sample sizes per group did not justify the use of parametric tests either. Therefore, we selected the Mann-Whitney U-test for the analyses. In the case of Affinity, a statistically significant difference was found when showing a conventional video and not telling about AI ($M = 3.16$, $SD = 1.11$) and when showing the same video claiming it to be modified by using AI ($M = 4.15$, $SD = 1.22$) ($p = 0.02$). A parallel outcome was found when comparing whether or not the disclaimer affected Warmth: the person was rated to be warmer when having an AI disclaimer ($M = 5.07$, $SD = 1$) compared to when the disclaimer was not shown ($M = 4.27$, $SD = 1.25$) ($p = 0.04$).

Another evaluation was implemented by recruiting 328 MTurk testers to evaluate the same videos by using the same scales, but allowing the use of ANOVA in the analysis. For that data, no statistically significant differences were found when comparing pairwisely

whether or not the AI-disclaimer had an effect on how conventional video, AI-created video, and AI-modified video were perceived. Consequently, we compared the videos leaving AI disclaimer from the equation (Figure 2). This comparison revealed statistically significant differences in all features between the conventional video and the others ($p < 0.01$). The only statistically significant differences between AI-modified and AI-generated videos were in humanness ($p < 0.01$). Results of these evaluations show that telling users about using AI has only a small but statistically significant difference in how users perceive the person on the video. Comparisons show that conventional videos outperform AI-manipulated and AI-generated ones, with little difference between the latter two.

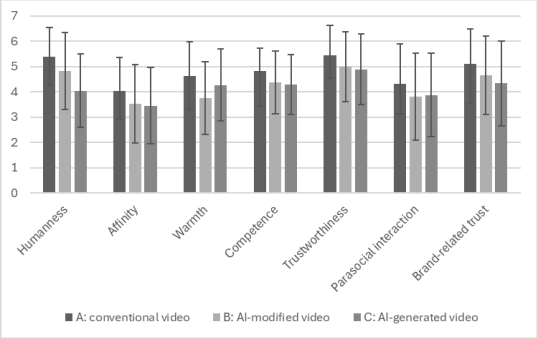


Figure 2: Evaluation results of echelon E8

In E11 (VC: accuracy), language experts focused on the domain-specific accuracy of the translated videos. From three to four iterations were needed in order to have a polished output.

Four instances of user evaluation E17 (VC: utility, usability, accuracy) were concurrently performed after the system design concentrated on an LLM-driven interactive text clone. Due to the system's nature as an LLM-powered, constantly evolving expert system, recruiting single test users was deemed relevant on a weekly basis. Four users were challenged to also compare the answers they get for exactly the same question or task from the generic ChatGPT version that they use and the proprietary expert fine-tuned GPT text clone, alternating between whom they ask first. In sum, these individual test users, representing different businesses (event organization, private wealth management, SaaS business and IT services), reported that their acceptance and perceived usefulness of the system increased significantly when the AI was introduced as an extension of a real expert, particularly in personalized or pedagogical contexts, highlighting the role of human framing and authorship in shaping trust. Compared to generic large language models, the text clone was perceived as more concise, context-relevant, and pedagogically aligned, and more to the point, especially when used alongside video-based learning, suggesting value-added complementarity between static and interactive learning modalities. A recurring limitation identified by users was thematic repetition, with the system tending to over-rely on a narrow subset of familiar frameworks, such as blue ocean strategy and value-based selling, indicating the need for a larger knowledge corpus. All respondents proposed enhancements to system transparency, including timeline-based documentation of the training corpus and the ability to track or personalize response origins, emphasizing a growing user expectation for epistemic traceability in AI systems. The notion of the 'digital-me' emerged strongly from user reflection, with participants noting both increased emotional resonance and social acceptance when AI outputs were mediated through human interaction.

As the final engagement at the end of the ^{e}DSR process, the expert system was observed in use with young entrepreneurs seeking advice for their startups, who ultimately inspired the vision later coined as the “Omnichannel Oracle” - an immortal digital-me expert system. With their lenses, they viewed the expert system as a concept of an "immortal expert persona", encapsulating the vision of a persistent, evolving AI entity that retains and refines its knowledge over time.

5. Discussion

5.1. Theoretical Advancement of the Digital-Me and Its Acceptance

One of the key findings to emerge from the development and deployment of the digital-me system concerns the mechanisms by which users interpret and accept AI-mediated representations of real individuals. Across demonstrations, internal trials, and user evaluations, it became apparent that acceptance was driven less by the technical fidelity of the system and more by the perceived social proximity between the avatar and its human counterpart. When AI-generated outputs were distributed through channels associated with the real expert—such as personal social media, direct communication with familiar audiences, or with an explicit endorsement—the responses were notably more positive. By contrast, the same content, when delivered anonymously or in impersonal settings, was often received with skepticism. These observations suggest a shift away from traditional constructs, such as visual or vocal realism (Epley et al., 2007; Wang et al., 2015), toward a new explanatory variable: real-human

closeness. This construct captures the degree to which the system is perceived to be contextually and relationally tied to a specific individual, regardless of its representational accuracy.

This insight aligns with recent studies highlighting the limited explanatory power of the uncanny valley effect in modern AI interactions. Miller et al. (2023) demonstrate that users often rate AI-generated faces as more convincing than real ones, suggesting that perceptual realism alone does not determine trust or authenticity. Similarly, Ahmed & Chua (2023) show that context plays a decisive role in deepfake interpretation, particularly when the viewer has prior familiarity with the subject. In our study, the effectiveness of the digital-me was not tied to how human-like the system appeared, but rather to the clarity of its human origin and social anchoring. This suggests that future models of avatar and AI-agent acceptance should account not only for perceptual features but also for the social signals that indicate authorship, ownership, and legitimacy.

A recurring theme across feedback was the importance of human control and explicit disclosure. Systems in which the expert was visibly or procedurally involved were more likely to be trusted, even if the outputs themselves were AI-generated. This finding resonates with earlier work on source credibility and parasocial interaction (Ohanian, 1990; Rubin et al., 1985; Hartmann & Goldhoorn, 2011) but introduces an important nuance. It was not merely the appearance of a person that fostered trust, but the perception that the person had authorized, influenced, or reviewed the output. Disclosure, when combined with indicators of human consent and control, enhanced legitimacy more than disclosure alone. In this regard, the communication of AI involvement should not be treated as a binary condition, but as a layered narrative about human agency, curation, and accountability.

These findings also challenge the underlying metaphor of the AI "agent" as an autonomous entity. The digital-me prototype points instead toward a cyborg-like configuration, where the system functions not as an independent agent but as an extension of the human expert. This reframing aligns with a growing body of research that considers AI as augmentation rather than substitution (Chen et al., 2024). The system is most effective when it retains a traceable connection to the human's thinking, tone, and intent. As such, the objective is not to simulate an artificial identity but to scaffold and project the expert's cognitive and communicative capacities into new contexts. This perspective has implications for design strategies, calling for tighter integration between the individual and the system across both technical and communicative layers.

A more speculative, yet increasingly relevant insight relates to the temporal scope of identity preservation. In later phases of the project, participants began to reflect on the system's potential to serve as a form of digital continuity, or even immortality. With sufficient source material and proper consent, it appears feasible to reconstruct not only an individual's expertise but also their narrative voice, rhetorical style, and characteristic reasoning patterns. While the idea of posthumous avatars may have once seemed hypothetical, it now raises genuine ethical and legal concerns. Should individuals, for instance, specify in their wills whether their virtual self can be maintained after death? Might future legal frameworks include a "digital self-determination clause," alongside living wills and organ donation consent?

Together, these observations support a theoretical extension beyond current AI acceptance models. We propose that personalized AI systems require a new conceptual framework grounded in three interrelated dimensions: perceived human control and authorship, social and relational closeness to the real individual, and temporal continuity across life stages. These elements constitute a basis for a Digital-Me Acceptance Model, distinct from traditional chatbot or avatar models, and more attuned to systems that are socially and epistemically tied to named individuals. Future research should explore the governance, ethics, and long-term societal implications of such persistent digital selves.

5.2. Changes to ^{e}DSR thinking in AI-powered system development

Continuously changing system performance. Although artificial intelligence technologies were integrated from the beginning, a fundamental transformation in the nature of development was observed once system operation, rather than only the creation of discrete artifacts, became dependent on AI functionality. At this point, the process no longer resembled a sequence of bounded engineering tasks. Instead, it shifted into the continuous management of a system whose core behavior was influenced by statistical inference, adaptive learning, and external change. The system's performance proved inherently unstable. Within sessions, the language model adapts to user input, showing short-term learning beyond rule-based programming. Between sessions, responses varied despite consistent prompts. Meanwhile, backend model updates—such as infrastructure changes, version

migrations, and silent fine-tuning—introduced further shifts in behavior. These changes, outside the development team's control, directly impacted consistency and user experience. Such dynamics challenge core assumptions in traditional DSR, which presuppose artifact stability between design and evaluation phases (Hevner et al., 2004; Peffers et al., 2007). When system behavior evolves independently, these assumptions collapse. The eDSR methodology (Tuunanen et al., 2024) offers a more suitable alternative by accommodating asynchronous, recursive change and shifting performance baselines. As Annepaka & Pakray (2025) note, large language models behave as probabilistic systems shaped not only by prompt engineering but also by opaque and evolving backend processes.

Changes in the nature of system development work. The nature of development shifted markedly as the system moved from static feature creation to active AI-based operation. Iteration accelerated, with interventions occurring too frequently and fluidly to be distinguished as discrete phases. Stable versions became difficult to define, and the artifact increasingly resembled a moving target, shaped by prompt manipulation and external updates. This volatility led to more frequent movement between design stages. Development no longer followed a linear trajectory but returned repeatedly to earlier phases in response to behavioral drift or user feedback. A key enabler was prompt engineering, which allowed non-technical experts to directly adjust model outputs. This broadened participation but also introduced risk. In several instances, the system drifted toward vague or generic outputs, deviating from the expert voice it was designed to reflect. Reinforcing expert tone often led to excessive repetition and narrowness, revealing a core design tension between adaptability and identity preservation (Chen et al., 2024). Over time, development began to resemble operational supervision more than conventional system building. Activities focused on calibration, monitoring, and behavioral alignment rather than original construction. Like modern industrial automation, oversight became central to sustaining intended system performance.

Operational deployment as an important part of system design. Operational deployment did not mark the conclusion of development but became an extension of it. The system was deployed not once it was deemed complete, but when users were ready to adopt it in real-world conditions. This reframed deployment from endpoint to data-generating mechanism and raised the question of whether readiness should be defined by technical maturity or practical uptake. This logic aligns with Tuunanen et al. (2024), who emphasize that in complex, evolving systems, all echelons remain active throughout the design lifecycle. Deployment joins design, evaluation, and demonstration as recursive, interleaved activities. Real-world use generates insights that structured testing often fails to reveal. This development model mirrors practices in the gaming industry, where early releases serve to test market fit before refinement. Similar approaches are gaining traction in AI, where actual use reveals behavior patterns and edge cases that would otherwise remain invisible (Seymor et al., 2021). Rather than indicating completion, it acted as a mechanism for refinement. This recursive stance is central to the logic of eDSR, in which design knowledge is shaped not only by formal evaluation, but by continuous adaptation in live environments.

The rapid evolution of key system features highlights the importance of timing; as standard video outperformed AI-generated videos, focus was temporarily redirected toward the knowledge clone, with the option to integrate these approaches at the appropriate stage of technological development. Effective development increasingly depends not only on technical capability, but on the ability to orchestrate design decisions in sync with shifting technological conditions. Referring to the research question 3, staying attuned to the pace and direction of change has become a core requirement for maintaining relevance and capturing emerging opportunities.

6. Practical implications

Expert-controlled AI increases trust and adoption. The study suggests that acceptance of AI tools rises when users perceive the output as being under the explicit control of a named human expert. For example, users reported that they were more likely to trust the information when it came from someone they know, and that they preferred it when the tool reflected the expert's personality and tone of voice. Referring to research question 1, this implies that organizations introducing AI-based knowledge tools should invest in branding them through recognizable and credible individuals, rather than relying on anonymous bots.

Contextual deployment enhances perceived value. Users engaged more deeply with the system when it was integrated into educational activities, course materials, or professional workflows. Respondents note that the clone works best when used in conjunction with video training.

Human framing remains essential. Despite the sophistication of the underlying language models, many users emphasized the importance of a human

presence or content endorsed by a reliable expert. This reinforces that AI-based learning or support systems should not aim to replace expert presence but rather extend it.

Redundancy and content saturation reduce engagement. Even when technically accurate, repetitive answers led to a decline in enthusiasm. Partially answering to the research question 2, this suggests that cloning expert knowledge should include mechanisms to avoid overusing the same concepts and to diversify content responses over time.

Disclosures and control boost legitimacy. Clarity on the source of information was frequently requested. Users requested real references and expressed a need for greater transparency regarding the content of the expert knowledge repository. Including visible disclaimers and documentation of the expert knowledge base may increase credibility and foster better learning outcomes.

7. Limitations and future research

This study relied on a static knowledge repository, which ensured clarity but did not reflect the evolving nature of expert knowledge. A continuously updated knowledge base would likely accelerate design iterations and blur the boundaries between system design, user interaction, and ongoing development.

The evaluation context also presents limitations. The small number of demonstration participants (e.g., for voice cloning) and the reliance on student samples constrain the generalizability of the findings. The analysis method in E7 was revised to enable more detailed evaluation with improved data rather than to compare datasets. A larger sample size reduced the risk of type I errors, yielding p-values below 0.01 or 0.001 when significant differences were observed. While these groups provided valuable initial insights, future studies should expand to more diverse populations and real-world contexts to validate and extend the results. Moreover, this study did not directly address ethical, privacy, or consent-related implications, which remain essential considerations for broader deployment.

Future studies should also investigate systems that incorporate real-time feedback and knowledge updates to simulate more realistic deployment contexts. One particularly promising avenue is the development of a Digital-Me Acceptance Model, building on established technology acceptance frameworks (Venkatesh & Davis, 2000). This new model could include dimensions such as perceived proximity to a real human, continuity of identity, user control, and the symbolic meaning of interacting with one's digital self, especially in educational and expert-based applications (see also Venkatesh & Davis, 2000; Nass & Moon, 2000; Nowak & Biocca, 2003). Longitudinal testing of different versions of text clones using standardized survey instruments would also allow researchers to track changes in user attitudes, content perception, and trust over time. In the longer term, the development of such systems appears to converge toward highly personalized, continuously learning constructs that integrate three essential layers. These include a private expert-level knowledge base, real-time access to broader public information, and a socially recognizable persona that expresses the individual's style, humor, tone, and domain-specific ways of thinking. These agents could eventually operate across all communicative interfaces, ranging from text and video to embodied interaction. This developmental trajectory also raises novel societal questions. Will individuals maintain rights over the continuation of their digital selves after death, similar to living wills or donor registries? Will different versions of the digital-me be created for different domains of life, resembling the domain-specific deities of polytheistic traditions, rather than converging into a single, all-encompassing personal AI? Further investigation is needed in avatar-based human-computer interaction, particularly in areas such as mirroring nonverbal behavior and providing emotionally aligned responses (Reeves & Nass, 1996). Another important direction is linguistic personalization. As systems begin to adopt users' own communication patterns, including vocabulary, tone, and stylistic choices, it is worth studying how deeply such adaptation should extend and under what ethical constraints (van Dijck, 2013). This study provides a foundational contribution to the theorization of digital expert identity and lays the groundwork for a broader interdisciplinary agenda that connects design science, human-computer interaction, media studies, and digital ethics.

8. References

Golovianko, M., Gryshko, S., Terziyan, V., & Tuunanen, T. (2023). Responsible cognitive digital clones as decision-makers: a design science research study. European Journal of Information Systems, 32(5), 879-901.

Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W. X., Wei, Z., & Wen, J. (2024). A survey on large language model based autonomous agents. Frontiers of Computer Science, 18(6), Article 186345.

Agiollo, A., Cavalcante Siebert, L., Murukannaiah, P. K., & Omicini, A. (2024). From large language models to small logic programs: Building global explanations from

disagreeing local post-hoc explainers. Autonomous Agents and Multi-Agent Systems, 38(2), Article 32.

Annepaka, Y., & Pakray, P. (2025). Large language models: A survey of their development, capabilities, and applications. Knowledge and Information Systems, 67, 2967–3022.

Chen, Jin; Liu, Zheng; Huang, Xu; Wu, Chenwang; Liu, Qi; Jiang, Gangwei; Pu, Yuanhao; Lei, Yuxuan; Chen, Xiaolong; Wang, Xingmei; Zheng, Kai; Lian, Defu; Chen, Enhong (2024). When large language models meet personalization: perspectives of challenges and opportunities. World wide web, Vol.27(4)

Tuunanen, T., Winter, R., & vom Brocke, J. (2024). Dealing with complexity in design science research: A methodology using design echelons. MIS Quarterly, 48(2), 427–458.

Hevner, A. R. (2007). A three cycle view of design science research. Scandinavian Journal of Information Systems, 19(2), 87-92.

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.

Liang, S., Wei, Z. (2024). Understanding Users' App-Switching Behavior During the Mobile Search: An Empirical Study from the Perspective of Push–Pull–Mooring Framework. Behav. Sci. 2024, 14(11), 989

Mcuba, M., Singh, A., Ikuesan, R. A., & Venter, H. (2023). The effect of deep learning methods on deepfake audio detection for digital investigation. Procedia Computer Science, 219, 211–219.

Miller, E. J., Steward, B. A., Smith, J. K., Johnson, L. M., Chen, Y., Patel, R., & Dawel, A. (2023). AI hyperrealism: Why AI faces are perceived as more real than human ones. Psychological Science, 34(12).

Fiedler, A., & Döpke, J. (2025). Do humans identify AI-generated text better than machines? Evidence based on excerpts from German theses. International Review of Economics Education, 49, 100321. Elsevier.

Seymour, M., Lingyao, Y., Dennis, A. R., Riemer, K. (2021). Have We Crossed the Uncanny Valley? Understanding Affinity, Trustworthiness, and Preference for Realistic Digital Humans in Immersive Environments. Journal of the Association for Information Systems (2021) 22(3), 591-617

Köbis, N., Doležalová, B. & Soraperra, I. (2021). Fooled twice: People cannot detect deepfakes but think they can. iScience, Volume 24, Issue 11.

Ahmed, S. & Chua, H. W. (2023). Perception and deception: Exploring individual responses to deepfakes across different modalities. Heliyon, Volume 9, Issue 10.

Zhou, X., Kim, S., & Wang, L. (2019). Money helps when money feels: Money anthropomorphism increases charitable giving. Journal of Consumer Research, 45(5), 953–972.

Aaker, J. L. (1997). Dimensios of Brand Personality. Journal of Marketing Research, vol XXXIV (August 1997), 347-356.

Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. Journal of Advertising, 19(3), 39-52.

Yuan, L. & Dennis, Al. (2019). Acting Like Humans? Anthropomorphism and Consumer's Willingness to Pay in Electronic Commerce. Journal of Management Information Systems Volume 36, 2019 - Issue 2.

Hu, Q., Lu, Y., Pan, Z., Gong, Y., & Yang, Z. (2021). Can AI artifacts influence human cognition? The effects of artificial autonomy in intelligent personal assistants. International Journal of Information Management, 56, 102250.

Rubin, A. M., Perse, E. M., & Powell, R. A. (1985). Loneliness, parasocial interaction, and local television news viewing. Human Communication Research,12, 155–180

Judd, C. M., James‐Hawkins, L., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: understanding the relations between judgments of competence and warmth. Journal of Personality and Social Psychology, 89(6), 899–913.

Epley, N., Waytz, A. & Cacioppo, JT. (2007). On seeing human: A three-factor theory of anthropomorphism. Psychological Review. 2007;114:864–886.

Hartmann, T., Goldhoorn, C. (2011). Horton and Wohl Revisited: Exploring Viewers' Experience of Parasocial Interaction. Journal of Communication 61(6):1104 – 1121.

Ferstl, Y., Thomas, S., Guiard, C., Ennis, C. & McDonnell, R. (2021). Human or Robot?: Investigating voice, appearance and gesture motion realism of conversational social agents. 76-83.

Higgins, D., Zibrek, K., Cabral, J., Egan, D. & McDonnell, R. (2022). Sympathy for the digital: Influence of synthetic voice on affinity, social presence and empathy for photorealistic virtual humans. Computers & Graphics.

Wang, S., Lilienfeld, S. O., & Rochat, P. (2015). The uncanny valley: Existence and explanations. Review of General Psychology, 19(4), 393.

Jia, R., Steelman, Z. R., & Reich, B. H. (2017). Using Mechanical Turk data in IS research: risks, rewards, and recommendations. Communications of the Association for Information Systems, 41, 301-318.

Patrizi, M., Šerić, M. & Vernuccio, M. (2023). Hey Google, I trust you! The consequences of brand anthropomorphism in voice-based artificial intelligence contexts, Journal of Retailing and Consumer Services, Volume 77

Komiak, S. & Benbasat, I. (2006). The effects of personalization and familiarity on trust and adoption of recommendation agents. MIS Quarterly Vol. 30 No. 4 pp. 941-960/December 2006.

Stein J.-P., Messingschlager T., Gnambs T., Hutmacher F., Appel M. (2024). Attitudes towards AI: measurement and associations with personality. Scientific reports, 14 (1), pp. 2909.

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Management Science, 46(2), 186–204.