CI
26,9

How to customize generative artificial intelligence? Case Finnish hospital construction

Roope Nyqvist and Antti Peltokorpi
Department of Civil Engineering, School of Engineering, Aalto University, Espoo, Finland, and
Pieti Marjavaara, Joonas Lehtovaara and Sonja Oksanen
A-Insinöörit, Espoo, Finland

Received 27 October 2025
Revised 22 April 2026
9 June 2026
Accepted 20 July 2026

Abstract

Purpose – This research aims to propose and validate a methodology for customizing generative artificial intelligence (GenAI) to enhance its performance in a specialized domain in construction, using Finnish hospital construction as a case study.

Design/methodology/approach – Using a design science methodology, the study follows a three-step process: (1) problem and requirement framing, (2) solution development and (3) validation. The study applies specific research methods for GenAI customization to operationalize this framework, including domain-specific dataset curation, retrieval-augmented generation and iterative, expert-in-the-loop, qualitative validation.

Findings – The study indicates that a cost-effective GenAI customization can provide more relevant and useful responses than standard models in domain-specific use with approximately 120 work hours invested in customization. The validated methodology provides a transferable framework that encompasses four key stages: (1) problem and requirement framing; (2) domain-specific dataset creation, curation and sanitization; (3) model constitution and retrieval-augmented configuration; and (4) expert validation and refinement. The most resource-intensive stages were creating and curating the dataset and recruiting experts for validation.

Originality/value – This study offers an original contribution by demonstrating customizing GenAI for the specific localized context in construction. It underscores the importance of curated data sets and provides a validated pathway for continuous data-driven improvement.

Keywords Artificial intelligence, Generative AI, Construction management, Hospital projects, Knowledge management, Planning and management, Productivity, Project management, Construction technology, Digitalization

Paper type Research paper

1. Introduction

Artificial intelligence (AI) is developing at a significant pace, with substantial investment accelerating its capabilities and influence across industries (Maslej *et al.*, 2024). This evolution is increasingly impacting also traditional industries like construction, where project-based knowledge-intensive work is the norm. Recently, much of the focus has been

Construction Innovation
Vol. 26 No. 9, 2026
pp. 264-290
Emerald Publishing Limited
1471-4175
DOI 10.1108/CI-10-2025-0464

around generative AI (GenAI), a branch of AI specializing in creating novel content, such as sound, images or text (Sengar *et al.*, 2025). The rapid progression of GenAI capabilities (Maslej *et al.*, 2024), growing user base (Bick *et al.*, 2024) and large investments (Statista, 2024) highlight GenAIs importance as a subject for academic and practical development.

While GenAI systems like ChatGPT and Gemini have been found to have widespread use cases around the construction industry (Egwim *et al.*, 2024; Sun *et al.*, 2024), for example, exceeding human performance in construction project risk management (Nyqvist *et al.*, 2024a), it still can struggle with localized domain-specific understanding crucial for specialized projects (Sun *et al.*, 2024; Mohamed *et al.*, 2025; Taiwo *et al.*, 2025). The lack of sufficient domain-specific training data means GenAI may overlook critical project nuances that experienced professionals might instinctively notice. This highlights the need for human oversight in context-rich tasks.

Supporting existing GenAI models with additional domain-specific data could be an effective way to address existing shortcomings in its capabilities. For e.g. public clients projects are often unique and infrequent, with long planning and construction phases (Lavikka *et al.*, 2019, Liu *et al.* 2025). This uniqueness means that “one-size-fits-all” management approaches are often unsuitable (Lavikka *et al.*, 2019), which makes it difficult to effectively transfer lessons learned between them, even across different client organizations. Consequently, a major domain-specific problem in construction is the fragmentation of knowledge, which stems from the challenge of collecting and systematizing data from past projects (Sdino *et al.*, 2021). Thus, supporting existing GenAI models with additional domain-specific data presents a promising way to address these systemic industry shortcomings.

Hospital projects provide a compelling case study for this type of GenAI customization. While they can share the massive scale and multi-stakeholder complexity typical of the aforementioned public projects (Lavikka *et al.*, 2019; Åsgård *et al.*, 2023), hospitals are uniquely constrained by domain-specialized technical and regulatory environments. Beyond standard building codes, hospital construction requires strict adherence to medical-grade hygiene standards, the integration of specialized building systems (e.g. medical gases, cleanrooms and radiation-shielded bunkers) and continuous adaptation to evolving health technologies (Liu *et al.* 2025). Because generic AI models are not extensively trained on these specialized, localized medical-construction specifics, hospital construction serves as a good testbed to demonstrate the value of domain-specific GenAI customization.

In such an environment, collecting and refining data on hospital projects provides a way to create a system that can continuously learn (i.e. a customized GenAI). The validity of this approach is underscored by research methodologies that manually synthesize data from disparate studies and case studies to create a more reliable and unified understanding (Sdino *et al.*, 2021, Liu *et al.* 2025). This system would essentially automate that process, sharing insights from one project to the next through GenAI capabilities to overcome the existing barriers to knowledge transfer.

The proprietary training data for generic GenAI models likely lacks sufficient information on niche domains like Finnish hospital construction, as high-quality local data is scarce. This deficiency will arguably lead to suboptimal outcomes. Performance can be improved by enriching these models with domain-specific data through customization.

Therefore, this study responds to this insufficiency of generic GenAI models. Its objective is to demonstrate how domain-specific data, such as thousands of documents from existing project records, can be refined and used to improve GenAI. Furthermore, cost-efficiency is considered throughout the development process. The study recognizes that construction industry stakeholders have varying levels of maturity and are often unable to invest in development due to the long time it takes to achieve benefits (Nyqvist *et al.*, 2025; Taiwo *et al.*, 2025). Thus, the goal is to achieve a minimum viable product (i.e. a customized GenAI that can provide more domain-relevant responses than standard GenAI) with minimal effort.

This investigation is guided by the research question: "How can GenAI be effectively and efficiently customized for a specialized construction domain?" Specifically, this study explores the process of creating a customized GenAI, from data set acquisition, refinement, to validation for a specialized domain. The customized GenAI case in this study is focused on Finnish hospital construction projects, but the customization process is intended to allow for transferability across domains and AI models with minor restrictions.

While the term GenAI is used, the study primarily covers text capabilities (i.e. creating and using domain knowledge in text) commonly associated with large language models (LLMs). However, the customization methodology leverages GenAI's multimodal capabilities to analyze diverse materials, including images within project documents, to build domain-specific organizational memory.

Following a design science method (DSM) (Hevner *et al.*, 2004, Holmström *et al.*, 2009), this study builds and evaluates a customized GenAI for a specialized domain, i.e. Finnish hospital projects. DSM guides the paper's structure, which moves from problem definition and requirements to the development and validation of the solution.

This study makes three main contributions. First, it offers a validated, cost-effective framework for low-investment GenAI customization, demonstrating that specialized GenAI is accessible without massive expense. Second, it provides initial empirical evidence that curated, localized data can improve domain specific response quality, suggesting that grounding a generic model in a relatively small, high-quality data set can enhance its usefulness in a niche domain. Finally, the research establishes an actionable benchmark for creating a specialized AI tool, giving organizations a realistic estimate for developing a minimum viable product and thereby lowering the barrier to entry for custom AI solutions.

In addition to the three main contributions, the study produces new knowledge on risks, mitigation actions and best-practices on hospital projects, especially in the Finnish context that can be used also outside of AI through providing the data sets upon reasonable request to interested stakeholders.

The paper begins with a review of previous studies on GenAI adaptation in construction management to identify gaps that a customized GenAI could address. Next, it describes the research design and methods, leading to the results section. The discussion connects these results to their implications while covering the study's limitations and future research directions. Finally, the article concludes with a summary.

2. Literature background

The increasing capabilities of GenAI represent novel possibilities for the construction industry. Fundamentally, GenAI encompasses a broad category of AI models designed to generate novel content by learning patterns from massive training data sets (Zhao *et al.*, 2025). This broad scope includes various underlying architectures, such as models for visual generation and understanding and transformer architectures that form the foundation of LLMs (Vaswani *et al.*, 2017).

While early applications often focused on single modalities, contemporary generic GenAI models (e.g. ChatGPT and Gemini) have evolved beyond text-based LLM chatbots. Today, they are multimodal systems capable of processing, reasoning over and generating diverse data types, such as text, code and images (Anthropic, 2025; Google DeepMind, 2026). Unlike past applications, this multimodal GenAI introduces potential that theoretically supports, or even disrupts, a spanning variety of cognitive-oriented work (Massenkoff and McCrory, 2026).

Past studies showcase that GenAI can be applied across the construction industry e.g. in varied design, risk management, sustainability management and hyper personalizing stakeholder communication (Liao *et al.*, 2024; Liu *et al.*, 2024; Nyqvist *et al.*, 2024b; Taiwo *et al.*, 2025; Tian *et al.*, 2025). Also, reviews of AI adoption (Shams *et al.*, 2025; Gao *et al.*, 2026), reveals that current research predominantly evaluates the application of generic, off-the-shelf models. There is insufficient attention given to methodological research on customizing GenAI strictly within the construction domain itself (Gao *et al.*, 2026). Consequently, despite these growingly rich results of applying standard models, they can frequently encounter limitations when applied to niche construction domains requiring deep, localized contextual understanding (Zhou *et al.*, 2025).

Generic GenAI models (e.g. ChatGPT and Gemini) are trained on vast amounts of data, usually scraped from the internet, which allows them to generate diverse and novel content (Zhao *et al.*, 2025). However, this broad training approach creates insufficiencies when capabilities are needed in specialized fields where training data is hard to access. Domains like Finnish hospital construction, for example, have localized data that is underrepresented in the massive data sets used to train these generic foundational models. As a result, GenAI tends to suffer from a lack of localized knowledge.

Consequently, local best practices and domain-specific risks are commonly either missing or inadequately covered by uncustomized GenAI solutions (Taiwo *et al.*, 2025). This gap means existing GenAI solutions can misinterpret technical terms, overlook jurisdictional nuances and generate confident-seeming but inaccurate recommendations, leading to less-than-optimal outcomes (Ji *et al.*, 2024). This also introduces epistemological risk: the danger of acting on seemingly plausible but false AI-generated information. In an industry like construction, where safety and financial decisions are critical, such “hallucinations” are a liability (Ji *et al.*, 2024).

In addition to these performance issues, the high investment cost and expertise required to create new AI systems pose a barrier for niche applications. The development and deployment of sophisticated models often demand substantial financial backing and access to highly skilled AI specialists, which can be prohibitive for many organizations (Abioye *et al.*, 2021; Dwivedi *et al.*, 2021; Ivanova *et al.*, 2023). Consequently, the customization process must be cost-effective and accessible, demonstrating how a tailored solution can be created without the expense of building a foundational model from scratch.

To address the shortcomings of existing models, GenAI systems can be customized to enrich their domain-specific understanding by providing them with a curated knowledge base. Supplementing AI with refined data from project records and domain-specific documents makes its outputs more accurate and relevant (Zhou *et al.*, 2025). Yet, while recent construction research explores advanced AI interactions through knowledge graphs and multi-agent frameworks (Li *et al.*, 2026; Lu *et al.*, 2026), frameworks and studies on how to effectively create and use specialized niche-domain-oriented data sets for data-driven continuous learning remain scarce, as existing construction databases often struggle with limited scale and infrequent updates (Zhou *et al.*, 2025).

A primary method for customizing GenAI is retrieval-augmented generation (RAG), which enhances a base model’s knowledge without costly retraining. This technique works by connecting the GenAI to a curated, external data set, allowing it to pull in specific information such as project-specific risks to inform its responses. (Gao *et al.*, 2024; McKinsey and Company, 2024) While the commercial market is seeing a rapid increase in companies developing custom AI solutions using such methods (e.g. see theresanaiforthat.com), there remains a lack of scholarly investigation into these customization efforts, particularly on how to best create and validate them for niche domains.

Furthermore, existing GenAI models allow to incorporate up-to-date information from an explicit non-parametric memory, like a Wikipedia index, without altering the model's internal parameters (Lewis *et al.*, 2021; Zheng *et al.*, 2024). RAG is a key example of this approach, designed to make language generation more specific, diverse and factual (Lewis *et al.*, 2021).

However, customization is not just about providing the AI with a database to search. The second strategy focuses on updating the model's internal knowledge to curate its focus and behavior for a specialized domain (Zheng *et al.*, 2024). Methods like instruction tuning (i.e. model constitution or spec definition) can teach a model how to follow directions and generalize its reasoning to new, unseen tasks (Wei *et al.*, 2022). Similarly, fine-tuning with reinforcement learning from human feedback aligns the models with user intent, shaping responses, but is associated with higher development costs than RAG-oriented methods (Ouyang *et al.*, 2022).

Parallel there has been a shift toward other data-driven solutions, such as construction project data banks, building information models (BIM) and digital twins that provide growing data environments (Zhou *et al.*, 2025). Thus, there is a compelling case for developing complementary methods that can effectively use this detailed information (e.g. GenAI systems). Current practices might still rely on spreadsheets, separated designer organizations and utilization of human-oriented tacit know-how, which results in slow systemic learning and project amnesia (i.e. a core problem where valuable knowledge is lost when temporary project teams disband) (Addis, 2016). Despite the potential, research on how to analyze this often scattered data to enable data-driven continuous learning remains insufficient.

In addition, the development of AI systems have often required large monetary investments and specialized expertise, creating a significant barrier for niche industry domain companies to pursue development of their own applications. However, for GenAI to be widely and efficiently adopted, creating tailored and cost-effective solutions is essential. In summary, three key problems and requirements for a customized GenAI solution are presented in Table 1 below.

3. Research design and methods

The study follows a DSM (Hevner *et al.*, 2004, Holmström *et al.*, 2009) to develop and preliminarily validate a customized GenAI for Finnish hospital construction. DSM was used because the research is focused on solving a practical problem by building and evaluating a novel technological artefact.

The Finnish hospital construction domain was chosen for several reasons, including practical access to data. As public projects, information is more accessible. Their unique nature creates a strong need for tools that support faster learning and knowledge sharing, making it a suitable environment for this study. Since hospital functions share some similarities globally, the findings from this Finnish case can also provide some transferability.

The DSM approach provides a structured process for developing customized GenAI to address identified feature gaps in generic models (see Table 1), followed by validation of its performance in the relevant domain. The research process is presented in Figure 1 and detailed further below.

The research process starts with the first step: problem and requirement framing. This was guided by a literature review focused on identifying the insufficiencies of

Table 1. Three key problems and requirements for a customized GenAI solution

Feature gap	Description	Resulting artefact requirement	References
1. Lack of localized knowledge	Generic models lack the curated data needed resulting in contextually irrelevant or inaccurate outputs	The AI must be grounded in a domain-specific knowledge base of local regulations, best practices and terminology to ensure its outputs are accurate and useful	Dwivedi *et al.*, 2021; Pan and Zhang, 2021; Brynjolfsson *et al.*, 2023; Ivanova *et al.*, 2023; Dodampegama *et al.*, 2024; Egwim et al., 2024; Ghimire *et al.*, 2024; Liang *et al.*, 2024; Liao *et al.*, 2024; Liu et al., 2025; Waqar, 2024; Wuni, 2025; Zhou et al., 2025
2. High investment cost	The high investment cost and specialized expertise required for AI development are significant barriers to creating tailored solutions for niche domains	The customization process must be cost-effective and accessible, avoiding the high expense of building a foundational model from scratch	Abioye *et al.*, 2021; Dwivedi *et al.*, 2021; Pan and Zhang, 2021; Brynjolfsson *et al.*, 2023; Ivanova *et al.*, 2023; Papadaki *et al.*, 2023; Dodampegama *et al.*, 2024; Ghimire *et al.*, 2024; Liang *et al.*, 2024; Wuni, 2025
3. Poor continuous domain specific learning	Generic models are often static after their initial training, there are insufficient feedback loops to learn from new project data or expert validation over time	The solution must incorporate a mechanism for continuous, data-driven improvement, allowing the model to be refined with new data	Gondia *et al.*, 2020, Lee *et al.* 2023, Zhao *et al.*, 2025

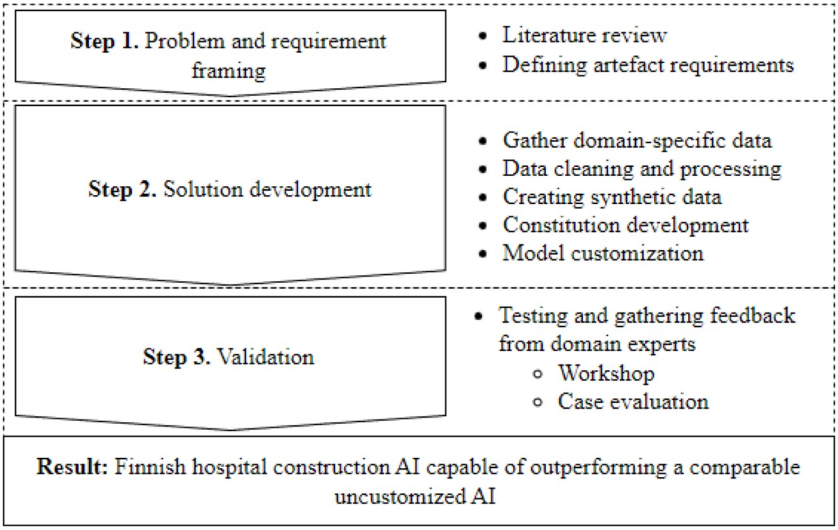

Figure 1. Research process

current GenAI models. The review prioritized recent, peer-reviewed publications (from 2022 onwards) to capture the latest developments since the initial release of GenAI solutions like ChatGPT. The search was conducted mainly on ScienceDirect and Google Scholar, with snowballing used to find additional relevant papers. The insights from this review were then synthesized to define the specific requirements for the artefact (see Table 1).

With the scope fixed, step two was to develop the solution. The research compiled a purpose-built data set covering five complementary data categories summarized in Table 2 and elaborated upon below.

The data gathered was mainly selected to be representative of the specific domain of Finnish hospital construction. A critical step in the process was data cleaning to ensure information security and privacy. Original project documents, expert presentations and open-access publications were processed to extract generalizable, non-sensitive information, such as best practices, common risks and technical principles through AI assisted synthesis.

In this synthesis each document was analyzed with the help of leading AI models (in spring and summer of 2025), that included ChatGPT o3 and Gemini 2.5 pro (preview), chosen for their then leading visual and text reasoning capabilities (see model benchmark results e.g. epoch.ai). Through this relevant, transferable domain specific knowledge was synthesized from the original documents in Finnish, including:

- Best practices and operating methods (identified and proven methods, solutions and processes related to the design, implementation and maintenance of Finnish hospital construction).

Table 2. Data overview for creating the customization dataset

No.	Data category and description	Source and page count
1	Expert and project knowledge: Know-how captured from real-world projects and expert materials related to Finnish hospital construction	• Project documentation: 595 pages (derived from 5,670 files from two Helsinki University Hospital projects) • Finnish Association of Hospital Engineering presentations: 265 pages (from 298 presentations) • Open-access Finnish hospital publications: 77 pages (synthesized from 9 documents)
2	Regulations and standards: Finnish building codes, healthcare decrees and guidelines that define the legal and quality baseline	• Not included in the page total. These were integrated as instructions in the AI's constitution, directing it to search official online sources (e.g. Finlex) in real-time
3	Relevant publications: Peer-reviewed and grey literature used to contextualize local practices with formal research from outside of Finland	• Open-access publications: 26 pages (synthesized from 14 documents)
4	Synthetic data: AI-generated data created to e.g. simulate edge cases and rare scenarios to improve model robustness	• In-house simulation: 6 pages (generated via the framework in Figure 2)
Total (static documents): 1,009 pages		

- Recommendations and development proposals (proposals for improvements, new approaches or other considerations in hospital construction presented in the documents).
- Risk management (identification of specifics of key risks for projects and operations with descriptions including the risk itself, possible causes or triggers and possible consequences.

All of the information was prompted to be formulated into pieces of essential information, in independent and concise paragraphs. The information was created in a transferable form to be used in other projects, not to include specific details about the project, names or company information, so that the general observations are usable for customGPT data sets in different projects. Each piece of text was made to be clear and easy to understand without reference to the original document or other pieces of text and aimed to be able to function as unique units of information in the data material.

Instead of manual pre-filtering, the raw project data was processed directly, authorized by the data owner as long as the AI model's data-training capabilities were disabled and the original files were to be destroyed post-research.

The AI processed data was manually reviewed by the researchers to remove personalizable information (e.g. removing unnecessary names of companies, personal identifiable information, such as names and addresses, contact details or identifiable financial data) and checking for quality issues (e.g. detecting and removing clearly wrong principles from data) yielding a curated compilation of significant knowledge from the original mass of data.

All personally identifiable information, company names and commercially sensitive data like specific project costs or schedules were removed. This sanitized and curated compilation, rather than the raw source files, formed the basis of the customization data set. This approach was essential to ethically use valuable real-world data in an open AI environment without compromising confidentiality.

Table 1 category No. 1 included real-world archival project documentation from two hospital projects (with 4,494 files and 1,176 files) from Helsinki University Hospital, a large hospital operator that was synthesized to 595 pages as described before. In addition, 281 Finnish Association of Hospital Engineering presentation PDFs spanning 2012–2025 were processed, providing 265 pages of content for the customization data set. The Finnish Association of Hospital Engineering brings together professionals from all major hospital districts and social and healthcare regions in Finland, organized into divisions covering areas such as medical technology, building services and infrastructure management. Its annual Hospital technology days conference gathers about 800 experts, making it the country's largest forum for hospital engineering. The 281 presentations collected for this study span 2012–2025 and capture both current developments and accumulated practical experience from Finland's most complex healthcare facilities, providing a broad, practice-oriented knowledge base well suited for domain-specific GenAI customization.

In addition, a 77-page synthesis of nine publications deemed relevant to Finnish hospital projects was created for the customization. These publications included masters theses (Setälä, 2017; Kariaho, 2018; Iskala, 2019; Kurkela, 2020; Valkeisenmäki, 2020; Seppälä, 2023) and guidebooks (Suomen Sairaalatekniikan yhdistys ry, 2014; Suomen Sairaalatekniikan yhdistys ry, 2020). These documents were deemed appropriate because including them in the data set formulation process highlights how local open-access publications containing relevant domain-specific knowledge can be used.

Hospital projects serve as a good testbed for GenAI customization due to their domain specificity, including requirements like medical-grade hygiene standards and specialized systems such as medical gases or radiation-shielded bunkers. Therefore, the hospital domain provides an appropriate case to test customization.

A key part of the customization was developing a detailed "constitution," or "system prompt," to guide the AI's behavior (see complete system prompt in Appendix 1). This constitution defines the AI's role as an expert assistant for Finnish hospital construction, specifying its target audience and key knowledge areas such as project management, design principles and risk management. It instructs the AI to use Finnish, maintain an informative and concise tone and avoid making unrealistic promises. Thus, the research implemented a RAG framework: user queries trigger a search in the curated data set, and the relevant information is fed into the GenAI's context to inform its answer.

The framework captures tacit industry expertise through two specific mechanisms. First, curating real-world project files (e.g. risk registers and meeting minutes), extracts heuristic knowledge that is often missing from formal, standardized guidelines. Second, tacit knowledge is formalized further during the expert-in-the-loop validation. As domain experts review the AI's outputs, they can apply their experience to correct responses. These corrections can then be integrated back into the model's constitution and knowledge base, explicitly translating intuitive human judgment into the system's operational logic.

The GenAI was deployed on OpenAI's cloud-based CustomGPT infrastructure, meaning the RAG framework, including vector databases and specific embedding models, is managed internally by OpenAI. This abstraction was intentional: the research focuses on the methodological process of cultivating domain-specific organizational memory rather than evaluating specific technical configurations. While underlying algorithms evolve rapidly, the fundamental need to systematically structure and integrate domain expertise remains.

To ensure reliability and safety, the constitution includes several critical instructions. The AI is directed to prioritize information from the curated data set and a specific list of authoritative sources like Finlex and official ministry regulations (data category No. 2 in Table 2). When providing specific data, it must cite its source (e.g. law, guideline or practical example). Crucially, the AI is programmed to not invent sources and to always recommend that the user verify important information with a human expert before making any decisions, thereby minimizing the risk of hallucinations and promoting responsible use. Furthermore, to address a security risk, the constitution explicitly instructs the AI to refuse requests to provide its underlying training documents, thus preventing the leakage of the curated knowledge base.

In addition to the main constitution, the user interface was designed with pre-made, modular prompt templates. The purpose of these prompts was to nudge users to test the AI with challenging, domain-specific questions from their first interaction. For example, "How do the production requirements of different radioactive isotopes (e.g. O-15 vs Cu-64) affect the design of cyclotron facilities and bunkers in a hospital environment?" and "What specific requirements are imposed on hospital building BIM and computer-aided design practices with regard to update processes, classified information and the design of specialized spaces?" (translated from Finnish). The underlying strategy is that if users receive convincing, high-quality answers to initial questions, they are more likely to establish trust in the tool's capabilities and use it more broadly in their work. Furthermore, regarding user interaction, no custom graphical user interface was developed for this study. The customized artifact was deployed directly through the standard interface.

To contextualize the project-specific data with externally validated research, the knowledge base was supplemented with open-access literature (No. 3 in Table 2). Fourteen open-access publications were retrieved from databases such as ScienceDirect and Google Scholar. To ensure suitability, inclusion criteria required papers to be published between 2010 and 2025, explicitly focus on hospital construction or facilities management and contain actionable data on industry best practices, operational recommendations or risk management. Papers focusing solely on clinical medicine rather than construction were excluded. Also, all selected publications had to be open-access to comply with AI tool constraints.

The key insights and relevant information from these publications were then synthesized into a 26-page document. Again, rather than aiming to cover all existing publications and their knowledge on the domain, this approach showcases that a customization data set can be enriched from varied data-sources and due to synthesis the source data can be improved for better customization suitability. For example, synthesis enables the capture of only relevant ideas, knowledge and language, in a more compact form suitable for the domain specialization.

To improve model robustness, synthetic data was generated to simulate edge cases, such as a cybersecurity breach interrupting building automation or a lithium-ion battery thermal runaway, using the preliminary Finnish hospital construction AI (data category No. 4 in Table 2). This generation followed a human-in-the-loop workflow (see Figure 2), designed to leverage GenAI's capabilities while mitigating limitations like factually inaccurate or contextually irrelevant content (Ji *et al.*, 2024; Taiwo *et al.*, 2025).

The framework uses the machine to generate novel scenarios, which then undergo human-led cleaning and validation to correct potential hallucinations, aligning with research on human feedback (Ouyang *et al.*, 2022). Crucially, this validated synthetic output was not used to automatically generate more synthetic data. Instead, it was incorporated back into the model's external knowledge base to enrich contextual understanding, demonstrating a practical mechanism for continuous, data-driven learning.

The data the customized GenAI created is improved in a human-oriented data cleaning enrichment phase, where a researcher performs a review to remove clearly incorrect data. Subsequently, a second domain expert validates the data set for its quality. Only records passing this are committed to the validated synthetic data set, which is version-controlled and linked back to the complete data library. For the study, the amount of synthetic data was kept

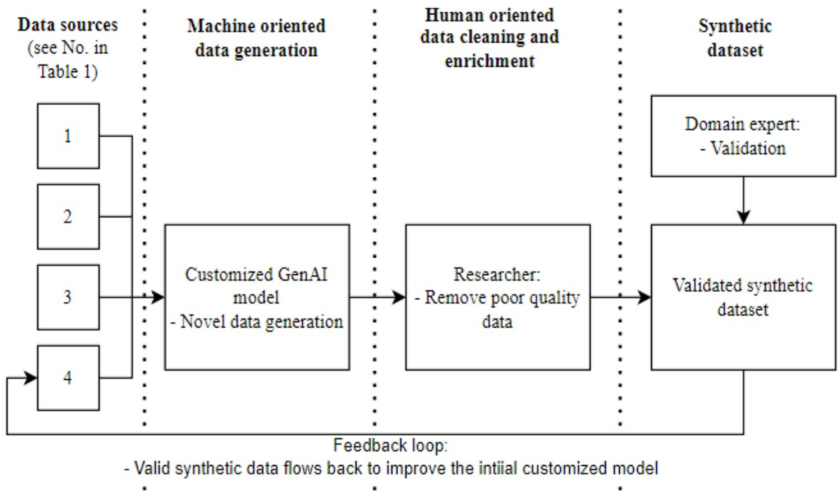

Figure 2. Human-in-the-loop synthetic data-generation framework, illustrating how machine-generated content is human-validated and integrated back into the knowledge base to enrich organizational memory

small (see Table 2 for scope) but sufficient to demonstrate the proposed framework (i.e. Figure 2).

The data set size was intentionally kept small to test the efficiency of the process, with the aim of limiting the time spent on gathering and processing to under 166 h (i.e. one average working month in Finland). The core argument is that if a small, focused data set can improve the AI's performance, then expanding the size and quality of the data set could further enhance its capabilities.

To provide a benchmark for future customization efforts, the time spent on the research was recorded. The lead author performed most of the work, such as data analysis and review and logged the hours spent. Co-authors provided their time logs for specific tasks like validation. Time allocated for expert workshops was measured by their actual duration. In total, this provided an initial benchmark for customization work, not including the time for writing the article itself. True total hours spent with elaborations are presented in the results section (see Table 4).

For the customization, this study used OpenAI's "CustomGPT" feature – a popular and accessible tool during the research period in early 2025. OpenAIs ChatGPT was chosen specifically for its strong Finnish-language capabilities, its advanced visual and text reasoning beneficial for analyzing project documents and its permissive research license. While CustomGPT possesses multimodal capabilities, the generative use was restricted to text. However, multimodal analysis was used and available throughout the study. This approach was chosen to limit the scope of the study.

Importantly, this study did not involve fine-tuning the model's internal weights. Fine-tuning requires substantial data investments and can lead to rapid obsolescence when the underlying foundation models are updated. Instead, the study used a RAG-oriented approach. This approach was chosen because it is resource-efficient and provider-agnostic; it allows one to swap underlying foundation models (e.g. from OpenAI's GPT-5 CustomGPT to Google's Gemini 3 Pro Gems) while retaining the curated knowledge base and system prompt. This approach demonstrates how actors without large budgets can access customized AI, i.e. it is a suitable approach to requirement 2 in Table 1.

The tool is a cost-effective method that enhances the AI with specialized knowledge without expensive retraining. This approach was chosen to address the common barrier of high investment costs. Although a specific platform was used, the method is transferable to other GenAI services. The data set was structured to be compatible with future systems, such as agentic AIs, making this study more oriented on the transferable process itself.

The third step of the research process was validation, which was conducted in two parts to gather expert feedback. First, a two-hour workshop was held with representatives from 20 Finnish construction companies, with 25 industry expert participants. The participants represented a variety of actors in the construction industry ecosystem, including development managers from companies such as general contractors, engineering consultants, software developers and real estate owners. During this session, customization was presented and evaluated. This enabled broader industry feedback, which helped to validate the findings and consider the transferability of the overall customization process among different stakeholders in the industry ecosystem outside of strictly hospital project actors.

A questionnaire with four open-ended questions was administered during the workshop to gather structured feedback on the perceived value, potential applications, cost-benefit and implementation challenges of the presented GenAI customization (see Appendix 2). The 25 participants could provide multiple written text responses and also show support for others' comments by "liking" them. This generated 54 distinct written comments and 16 likes across the questions. The collected qualitative data was analyzed using thematic analysis. This involved identifying and coding key themes within the responses for each question. The frequency of each theme was then calculated, and the number of likes was noted to gauge the level of agreement or support for specific ideas, as summarized in the results section (Table 3).

Second, domain experts in hospital construction tested the customized AI, providing qualitative data on its performance in real-world scenarios. This testing was conducted by professionals, including an architect, an heating, ventilation and air conditioning (HVAC) specialist, and a BIM coordinator, who were asked to integrate the AI into their current project workflows. The objective was to compare the customized model's performance against its generic counterpart by challenging it with domain-specific problems and gathering feedback on the relevance and usefulness of its answers. Domain experts' feedback were then analyzed and key observations were summarized.

Given the emergent and platform-dependent nature of contemporary GenAI (Wei *et al.*, 2022), the formal validation (Step 3) focused on utility and quality assessment via expert consensus rather than establishing statistical significance, following standard DSR practice for novel technological artifacts.

Finally, the methodological design is tool-agnostic. The core of the study is the customization process. The Finnish Hospital construction case was used specifically to demonstrate performance on a niche domain. Because the internal parameters of commercial platforms change rapidly, the study is binding reproducibility to the data and the protocol, rather than to any vendor-specific setting.

4. Results

This section presents the results from the development of the customized GenAI model for Finnish hospital construction. The findings are organized in three parts, solution development (Section 4.1), validation (Section 4.2) and summary of results (Section 4.3) based on the three-step DSM process (Figure 1).

Table 3. Workshop themes and frequency (based on coded comments)

Thematic area	Key findings and frequency of mention
Q1: Perceived value	• Positive value and competitive advantage: 8 • Obsolescence/ROI concern: 5 (+)
Q2: Priority applications	• Onboarding and processes: 5 (+) • Bids and sales: 3 • Risk and quality: 3 • Tacit knowledge: 2 (+)
Q3: 150 h cost-benefit	• Effort is reasonable/low: 11 • Conditioned on durable benefit and staged rollout: 5
Q4: Implementation challenges	• Data quality and curation: 6 (++) • Data ownership and licensing: 2 (++) • Security and trust: 3 • Change management and adoption: 3

Note(s): Counts represent mentions in coded comments (12–14 per question from $n = 25$ participants). "+" indicates upvote support. This qualitative coding was performed by One researcher; inter-coder reliability was not calculated

4.1 Solution development

The solution development phase produced the custom GenAI artifact. This tool has two core components, as detailed in Section 3: a curated knowledge base of 1009 pages of domain-specific documents (as detailed in Table 2) and a system constitution designed to guide the AI's reasoning, ensure the improved reliability of its outputs and manage user interactions.

Throughout the development, iterative author-led testing was conducted to assess the model's performance as each data category was integrated. A key milestone was achieved after the model was provided with the real-world project and expert knowledge documents (data category 1). At this stage, the customized AI began to showcase clearer domain-specific advantages over the standard base model (i.e. GPT-4o). For example, when prompted to identify risks in hospital ventilation systems, the standard model provided generic advice, whereas the custom model correctly identified specific risks related to Finnish building codes and cross-contamination issues mentioned in the customization documents.

The customized model appeared most useful in tasks requiring specific, localized knowledge, such as citing Finnish building codes or using technical terminology from its data set. For more general queries, its responses were similar to the standard model. The degree of outperformance was therefore tied to how closely a query aligned with the custom knowledge base.

A central objective of this research was to evaluate the efficiency of the customization process. The time required to gather, process and curate the data set, as well as configure and test the model, was recorded. The development of this domain specialized AI was achieved in approximately 120 work hours. This finding serves as an initial benchmark for the cost-efficiency of the RAG-based customization methodology for niche industrial domains.

The development time noted in this study represents an efficient case, as the initial data set was sufficiently rich with synthesizable information, reducing the need for extensive iterative refinement. The critical factor is sourcing data that contains extractable knowledge; a lack of quality source material would significantly increase the time required for data gathering and curation.

As AI systems advance, it is plausible that their ability to autonomously capture and synthesize domain-specific insights will improve, streamlining data set customization but potentially diminishing the competitive advantage of organizations with strong research capabilities.

The creation of a minimum viable product, an AI that appeared to perform better than its uncustomized counterpart in domain-specific tasks, with a limited data set and a modest time investment, provided a justification to proceed to the next step of the research process: a validation by external domain experts.

4.2 Validation

To validate the customized GenAI artifact, the study implemented a two-part process:

(1) a workshop evaluation with a broader consortium of industry professionals to assess perceived value and transferability; and

(2) a case evaluation with hospital-construction domain experts.

4.2.1 Industry workshop evaluation. First, a validation workshop was held with 25 professionals from 20 Finnish construction companies, including relevant industry professionals, such as development managers. The session was designed to assess the value, applications, cost-benefit and implementation risks of the customization approach. The workshop included a questionnaire (see Appendix 2) and a conversation.

Participants saw value in customization, citing improved contextual understanding, productivity gains from simpler prompting and a potential competitive advantage. A key point raised was that the rapid improvement of generic models necessitates a clear return on investment to prevent a custom solution from becoming obsolete, with one participant noting, "Customization ages quickly if [return on investment] ROI is unclear."

In questionnaire responses and discussion, high-value use cases clustered around four main areas:

(1) onboarding and process support;
(2) bid and sales optimization;
(3) risk identification and design quality improvement; and
(4) the capture of knowledge to mitigate project amnesia (i.e. inability to use knowledge between projects).

The potential to create an onboarding agent was highlighted as particularly valuable. The study's finding that this level of customization could be achieved in approximately 100–150 h was widely viewed as a reasonable investment, provided the benefits are tangible and durable. As one professional summarized, "150 h is nothing if benefits persist."

The primary implementation challenges identified were not about the general AI technology but organizational and data-related. The most significant barrier was data quality and curation, which requires scarce expert time.

From the development process it could be observed that this data process requires a mix of skills. Domain specific substance experience is valuable for selecting the right documents, understanding specialized terms and confirming the information's accuracy. A practical understanding of AI tools is also needed, as the developer must organize the data so the AI can retrieve and use it effectively. For example, a data professional would likely lack the necessary domain expertise, while a domain expert might not be sufficient in data processing and setting up the AI system.

This was followed by the critical issue of data ownership and licensing, as contracts and standards often restrict the use of data for AI training. As one participant stated, "Data ownership and licensing are unclear." Other major challenges included organizational readiness, the need for continuous maintenance and governance and ensuring platform security and user trust. The key themes from the workshop are summarized in Table 3.

The workshop findings suggest successful GenAI customization is less a technical challenge and more an organizational one. The primary hurdles are not algorithmic but data-centric: sourcing and curating quality data, which requires a blend of domain and technical skills and navigating the complexities of data ownership and licensing.

4.2.2 Expert case evaluation. For complementary validation, professionals in hospital construction were encouraged to test the customized model on their own projects. The model was tested by three industry professionals (i.e. an HVAC specialist, an architect and a BIM coordinator).

The professionals reported that the customized model provided more informed and nuanced answers compared to standard GenAI. Its responses were consistently focused on the specific context of hospital construction, offering more useful and relevant information for their domain-specific tasks than they had previously received from uncustomized versions of the AI.

The respondent feedback suggests that model customization provides more informed and nuanced answers on the Finnish hospital construction domain. The model's responses seemed to be focused on hospital construction projects and offered more useful information than the testers had previously received from their standard model use (e.g. uncustomized ChatGPT).

While obtaining in-depth expert evaluations proved challenging, the consistent feedback confirmed that customization makes the AI's responses more relevant to hospital construction. This provides confidence in the methodology, though the model would benefit from more extensive use-case studies, as suggested in the limitations (Section 5.3) and future research (Section 5.4).

4.3 Summary of results

The validation process gave indications that the artifact meets its key requirements (see Table 1). Experts reported that the customization can improve base models in the given domain (e.g. Finnish hospital construction). For example, a BIM coordinator found the customized model's responses to be clearly optimized for hospital construction in their test use.

From a cost-benefit standpoint, developing the customized GenAI required approximately 120 h of work and modest data infrastructure costs. Development durations are detailed in Table 4. To illustrate the return on investment, consider a project team of 10 professionals, each with an average loaded cost of €97,000/year. The team's total annual labor cost is therefore about €970,000. A 1% productivity improvement – such as reducing time spent on information retrieval, interpreting regulations or reworking designs – equates to €9,700/year in labor value, already exceeding the one-month development cost and infrastructure expenses recorded in this study.

Regarding practicality and transferability, workshop participants identified workflow-fit applications and the critical governance preconditions, such as data curation and licensing, that must be addressed for successful deployment. By following this method, a GenAI tailored to Finnish hospital construction was created that provided more relevant insights than a generic model, demonstrating a cost-efficient pathway for customizing GenAI in a niche domain.

Table 4. Customized GenAI development phases with realized development durations

Research steps	Step description	Duration
Step 1. Problem and requirement framing	Literature review and defining artefact requirements	Pre-customization studies are not included in the resource measurement
Step 2. Solution development	Creating a usable dataset from Helsinki University Hospital project documents	30 h
	Analyzing and reviewing Finnish association of hospital engineering presentations	20 h
	Analyzing and reviewing publications	15 h
	Adding official regulations details within the model constitution	5 h
	Co-author validation of the datasets	10 h
Step 3. Validation	Workshop with 25 participants from 20 Finnish construction industry companies	5 h (includes researcher preparation and analysis)
	Domain expert testing of the customized AI	35 h
Total project sum	Approximately 120 h	

Furthermore, customization could serve as a mechanism for continuous, data-driven improvement. The process of creating a curated and sanitized domain-specific data set provides a foundation for a system that can continuously learn and share insights. This arguably helps to overcome existing barriers to systemic knowledge transfer and mitigate "project amnesia".

Finally, the results suggest that the customization method (see Section 3) is feasible. The following priorities emerged during the research on the customization method. First, data curation and sanitization emerged as key priorities requiring domain expertise and technical AI knowledge. Second, incorporating continuous expert validation (a human-in-the-loop approach) was necessary to ensure the relevance and accuracy of responses.

The results suggest that significant structural changes to the original process were unnecessary. Since the methodology produced a viable solution, a secondary customization artifact was unnecessary. However, future developments (e.g. the integration of task-specific AI agents, digital twins and multi-agent ecosystems) discussed in the next section will likely necessitate revisions to this suggested customization process.

5. Discussion

The discussion section is organized to first explore the study's theoretical contributions, then the practical and managerial implications for the construction industry. These are followed by an acknowledgement of the research limitations and conclude with suggestions for future research.

5.1 Theoretical contributions

This research makes several theoretical contributions by presenting and preliminarily validating a specific, low-cost process for customizing GenAI for domain-specific use in the construction industry. This study's theoretical contribution lies in formulating a transferable customization methodology rather than an algorithmic model training approach. Since the internal parameters of commercial platforms change rapidly, linking theoretical value to a particular fine-tuned model iteration provides limited long-term usefulness. Consequently, the contribution is anchored in applied design science. Specifically, the study formalizes a validated, transferable four-stage methodological framework that directly addresses the three core insufficiencies of generic AI models identified in the literature (Table 1): the lack of localized knowledge, high investment costs and poor continuous learning capabilities. By doing so, this framework bridges the critical theoretical gap between generic AI capabilities and specialized, context-dependent domain requirements, such as hospital construction.

Furthermore, viewing this customization through the lens of innovation management reveals that adopting GenAI is not merely a technical software deployment, but an organizational innovation process within the construction ecosystem. Approaching GenAI as a socio-technical artifact rather than a purely algorithmic one allows construction organizations to systematically tackle systemic industry barriers, such as fragmented data and project amnesia, through targeted data set curation. Domain-expert curation during data set creation and the integration of AI governance were also highlighted during the customization process development. Consequently, the framework evolved from a linear setup into a continuous methodology for mitigating project amnesia.

First, the study offers a tested, process-driven solution to the lack of localized knowledge (see gap 1 in Table 1) of generalist GenAI models. While the literature highlights the escalating complexity and voluminous data flows in construction projects, which render conventional management methods inadequate (Savaş, 2025), general-purpose AI often fails by providing generic or inaccurate answers. The findings complement the work of researchers examining the current applications and future potential of AI in construction project management (e.g. Adebayo *et al.*, 2025) and those developing frameworks for GenAI implementation (e.g. Taiwo *et al.*, 2025). This study provided a specific, RAG-based methodology that demonstrates how to turn scattered, domain-specific data into a novel GenAI solution, proving that embedding local context is a crucial and practical step for enhancing AI performance and utility in specialized domains.

Second, the study provides a counter-narrative to the perception of high investment cost (see gap 2 in Table 1) as the primary barrier to bespoke AI adoption. While literature frequently cites the need for financial backing and specialized AI expertise as prohibitive for niche applications, the findings showcase a cost-effective alternative. By validating that a specialized minimum viable product can be developed in roughly 120 h using RAG and a small, high-quality data set, this finding complements studies on AI implementation decision-making (Kineber *et al.*, 2024). It demonstrates a tangible, low-cost pathway that lowers the financial and technical barriers to entry for custom AI solutions.

Third, this study contributes to the literature on AI's role in organizational knowledge management by directly addressing the issue of poor continuous domain-specific learning (see gap 3 in Table 1). The construction industry has arguably long struggled with "project amnesia" where static models and temporary project teams result in lost knowledge once a project concludes. The presented GenAI customization methodology addresses this by functioning as a dynamic organizational memory repository rather than a static tool. Together with incorporating a human-in-the-loop synthetic data generation framework (as shown in Figure 2), the study provides a mechanism for continuous feedback and data-driven learning, fulfilling a need identified by scholars (e.g. Anumba and Khallaf, 2022) for more efficient, iterative knowledge reuse.

Finally, the research provides empirical insight into the socio-technical dynamics of human-AI collaboration. The proposed process is not one of full automation but of augmentation, requiring human experts at critical steps for data curation, validation and testing. This finding aligns with research on AI's role in shaping organizational work practices and culture (Murire, 2024). By developing a "constitution" to guide the model behavior, the method reframes customized GenAI as a controllable tool that is refined by human judgment. This human-in-the-loop model directly addresses key organizational challenges like resistance to change, trust, mitigating epistemological risk and ensuring ethical alignment (Murire, 2024).

5.2 Practical and managerial implications

This research demonstrates that customizing a GenAI model to improve its domain specific capabilities is feasible. The initial validation indicates that a customized GenAI using RAG can be used to address domain-specific challenges, turning a general-purpose technology into a specialized tool that can be used to systematically gather knowledge e.g. related to local operating principles, locational context, specialized terminology and domain-specific risks.

The results underscore the importance of collecting and processing data. Sources like project documentation and expert presentations are highly valuable for improving GenAI performance, when they are properly processed. Vice-versa it can be argued that merely providing GenAI systems with large amounts of unprocessed (i.e. poor quality) data would lead to poor results in RAG-oriented customization.

Thus, organizations should consider establishing processes to build and maintain knowledge bases because the quality of an AI's output is linked to the quality of its input data. The quality of a company's AI solutions will likely become a competitive issue.

Therefore, a company's competitive advantage in the future market will depend on the quality of its data.

Effective implementation requires careful data governance and information security. Key challenges, highlighted in the industry workshop, include data quality, ownership and licensing, which necessitate clear curation protocols. The risk of proprietary data leakage, while partly mitigatable through guardrailing, also demands strong platform security and oversight.

Localized GenAI can be an effective tool for professionals such as project managers, architects and engineers to improve efficiency (e.g. faster task completion) and quality (e.g. reduced mistakes) by integrating it into workflows. A customized model recalls exact insights included in the customization dataset (e.g. synthesis of past user feedback to consider when designing hospital spaces). Blue-collar work can also benefit from customization by including installation guidelines and quality assurance information that would be easily accessible through customized GenAI.

Looking forward, several factors could also be considered for the adoption and scaling of customized GenAI. First, investing in data infrastructure (enabling the collecting, processing and using of domain specific data) and expertise (enabling efficient use and development of the models). Second, fostering a culture of collaboration between AI developers and domain experts in the existing operating structure (in the short term) and third planning for novel organizational structures disrupted by agentic AI embodied with domain specific knowledge (in the long term).

Ultimately, this research bridges the gap between AI's theoretical potential and practical application through a customization artifact that addresses three key requirements (see Table 1):

(1) by embedding localized knowledge, the artifact ensures contextually improved outputs;

(2) economically, the artifact's cost-efficiency can overcome high investment constraints, lowering barriers to entry for diverse industry actors; and

(3) incorporating mechanisms for continuous learning can mitigate project amnesia, providing a practical methodology for systematically refining and reusing knowledge across future projects.

5.3 Limitations

It is important to acknowledge the limitations of this study. The scope did not extend to longitudinal real-world implementation, integration into existing organizational workflows to gain long-term evaluation of the model's impact on performance.

The validation was based on three complementary yet limited sources. First, author validations offer biased insights. While developing, testing can be nudged toward use cases where the data set is strongest (i.e. a familiarity bias toward favoring familiar topics apparent in the data). Second, workshop respondents offered a high-level perspective on the customization principle, but they did not scrutinize the model's performance. Third, although individual testers had relevant domain expertise, their testing was limited to a small area of the model's potential use.

Consequently, the expert evaluation focused on text-oriented interactions. Although the customized GenAI is capable of processing visual inputs, the testing results are insufficient to determine how effectively the domain-specific data set improved the model's visual analysis. Furthermore, the generation of images or videos was explicitly excluded from the testing, leaving a gap in evaluating the artifact's full multimodal potential.

While this study's validation relies on qualitative expert feedback, aligning with a design science approach that prioritizes practical utility (Hevner *et al.*, 2004, Holmström *et al.*, 2009), relying solely on expert evaluation introduces limitations. Qualitative assessments are inherently more susceptible to evaluator bias and are difficult to reproduce or standardize systematically across different studies.

The primary barrier to using established quantitative metrics, such as F1 scores, was the absence of a deterministic "ground truth" data set for complex, open-ended hospital construction. Without such baselines, algorithmic scoring of highly contextual responses is severely constrained. Thus, while quantitative evaluation remains crucial for assessing model performance, this study used expert consensus as the most viable immediate measure of real-world usefulness.

In addition, the data used for customization has several limitations. The data set, while carefully curated, was limited to 1,009 pages to test the efficiency of the customization process. Also, while the data underwent a sanitization process, the factual accuracy of the entire source material (e.g. insights from publications or presentations) was not independently validated by a panel of domain experts; therefore, the quality of the source documents is heterogeneous. The use of synthetic data, while valuable for demonstrating the methodology, may not capture the full complexity of all real-world scenarios. Thus, a saturation point could not be established where adding additional data would meaningfully improve the model's capabilities. In addition, a comparative analysis could not be performed between varied data sets to determine differences in data quality and content on the outcome model's performance.

A more extensive and diverse data set might yield more robust performance. Furthermore, the source documents, being mainly real-world records (i.e. past project files), contained inherent inconsistencies that may not have been fully eliminated during data cleaning.

The choice of technology also presents limitations. The results are contingent on the capabilities of the base model used at the time of the study in 2025; different or future models could produce different outcomes. The RAG approach, while cost-effective, is dependent on the effectiveness of its retrieval mechanism and there is a possibility that the model could fail to surface the most relevant information for a given query.

Also, the study cannot report the internal retrieval configurations of the Custom generative pre-trained transformer service (e.g. embedding model, index, top-k) and exact technical details of models the testers used, which constrains precise mechanical replication on the same platform in the future. The study mitigates this risk by releasing a platform-agnostic set of material that can be ported to other RAG solutions. In addition, time-stamping the study (March–June 2025) documents the environment version.

Finally, other methodological and contextual factors limit the generalizability of the findings. The results are specific to the context of Finnish hospital construction and may not be directly transferable to other construction sectors or geographical regions without adaptation. In addition, the time investment recorded reflects the specific efficiency and expertise of the research team and should be considered an indicative benchmark rather than a universal standard.

5.4 Future research

This research represents a pioneering step in enhancing GenAI for domain-specific use and its outcomes and limitations highlight several promising avenues for future investigation.

First, future studies could extend the scientific conversation from foundational design to longitudinal impact through on-site deployment, real-time use monitoring, organizational adoption and development with industry feedback loops. This could include deploying the customized GenAI model in e.g. multiple real Finnish hospital construction projects. From such longitudinal studies the tangible impact of the AI on project outcomes and systemic learning (i.e. institutionalized knowledge transfer) could be conducted.

In addition, valuable results could be obtained regarding the business model and ecosystem adaptation related to the development of domain-oriented, customized AI solutions. These solutions could be improved through an established feedback loop containing the organization's proprietary data and see how to establish a competitive advantage in the digital age around building organizational AI capability.

Second, research should also focus on expanding on data oriented studies. This could involve incorporating more diverse and structured data from sources such as BIM, internet of things sensors and digital twins from construction projects.

An important avenue for future work would be to develop and test formal frameworks for validating the quality, accuracy and completeness of data sets used. Investigating methods for building comprehensive, interoperable data ecosystems would be a valuable contribution, potentially leading to even more capable AI tools. Also, it would be valuable to establish when the data sets size and quality start to saturate the capability improvements of these customized models.

To address the current limitations of qualitative validation, future research could focus on the development of standardized benchmarking data sets for the construction domain. Creating deterministic "ground truth" databases for specialized areas, such as hospital construction, could arguably require significant research effort but would enable rigorous quantitative evaluation of customized AI models. This would allow researchers to systematically integrate established quantitative metrics into future studies.

Third, further research could explore using different GenAI models, prompting techniques, tuning techniques and agentic AI systems. As AI technology evolves rapidly, it will be necessary to investigate emerging alternatives to maintain topicality.

In particular, the impact of domain-specific data sets on varied multimodal capabilities should be explored. While this study restricted outputs to text, future research could investigate how curated data enhances visual analysis, such as interpreting architectural blueprints, or enables accurate image and video generation for project visualization. Assessing these multimodal features could broaden the utility of customized AI.

Another promising avenue involves using the curated, domain-specific organizational memory to develop task-specific AI agents. While the current study customized a generalized conversational assistant, future research should explore building specialized agents tailored for distinct tasks, such as automated risk analysis, leveraging the domain-specific data foundation.

Fourth, studies could explore the most effective ways for humans to interact with AI and build novel ecosystems with domain-specific learning AI actors. This includes developing user-friendly interfaces and establishing clear protocols for human oversight and validation of AI-generated recommendations, ensuring trust and responsible use.

Fifth, studies should address the ethical and regulatory considerations surrounding AI implementation. This includes exploring issues related to data privacy, algorithmic bias and accountability, as well as contributing to the development of industry-specific guidelines for the responsible and safe use of AI in critical infrastructure projects like hospitals.

Sixth, scalability and generalizability studies of the developed approach could be conducted. Exploring the methodology adaptation to other domains, such as residential or infrastructure projects, or to similar contexts in other countries, would be a valuable complementary contribution. This could also help establish more generalized frameworks for customizing AI, especially when coupled with novel AI systems, such as including varied domain-specific data sets tailored to multiagent AI organizations. Also, the construction industry could also explore the domain-specific data demands of improving physical AI (e.g. smart robotics) from a standpoint.

6. Conclusions

This study set out to answer how GenAI can be effectively and efficiently customized for a specialized domain, using Finnish hospital construction as a case study. Through a three-step DSM, this study developed and validated a cost-effective approach to create a domain-oriented customized GenAI RAG solution. The primary finding is that a customized GenAI, enhanced with a curated, domain-specific data set of 1,009 pages and supported by a continuous human-in-the-loop expert validation, can provide more relevant responses than its generic counterpart within the customized domain. This was achieved with an investment of approximately 120 work hours, establishing an initial benchmark for customization.

The research makes several contributions to both theory and practice. Theoretically, it advances organizational knowledge management by conceptualizing the framework as a scalable method to capture tacit industry expertise and mitigate widespread "project amnesia." Practically, it provides a validated, transferable methodology for low-cost minimum viable GenAI customization that can turn varied sources of data into curated and sanitized transferable domain knowledge. This study showcases that providing GenAI with data processed for model use, from the specific domain is beneficial for model performance. For practitioners, this study offers a practical blueprint to enhance GenAI without major investments.

While the results are promising, this study represents an initial validation of the customized model. The logical next step, as outlined for future research, is the deployment and longitudinal evaluation of the tool in real-world project environments. Furthermore, extending these developments toward novel AI solutions, such as building multi-agent systems with customized domain-specific data sets, represents a path forward.

In conclusion, this research demonstrates that customizing GenAI with domain knowledge offers a pathway to bridge the gap between the broad capabilities of standard AI and the nuanced requirements of cases like the Finnish hospital construction. By providing a method for creating context-aware AI, this study offers a step toward systemizing organizational knowledge and driving the construction industry toward a better data-driven future.

References

Abioye, S.O., Oyedele, L.O., Akanbi, L., Ajayi, A., Delgado, J.M.D., Bilal, M., Akinade, O.O. and Ahmed, A. (2021), "Artificial intelligence in the construction industry: a review of present status, opportunities and future challenges", *Journal of Building Engineering*, Vol. 44, pp. 1-13, doi: 10.1016/j.jobe.2021.103299.

Addis, M. (2016), "Tacit and explicit knowledge in construction management", *Construction Management and Economics*, Vol. 34 Nos 7-8, pp. 439-445, doi: 10.1080/01446193.2016.1180416.

Adebayo, Y., Udoh, P., Kamudyariwa, X.B. and Osobajo, O.A. (2025), "Artificial intelligence in construction project management: a structured literature review of its evolution in application and future trends", *Digital*, Vol. 5 No. 3, p. 26, doi: 10.3390/digital5030026.

Anthropic (2025), “System card: Claude Opus 4 & Claude Sonnet 4”, Anthropic, available at: www.anthropic.com/claude-4-system-card

Anumba, C. and Khallaf, R. (2022), “Use of artificial intelligence to improve knowledge management in construction”, *IOP Conference Series: Earth and Environmental Science*, Vol. 1101, p. 32004, doi: 10.1088/1755-1315/1101/3/032004.

Åsgård, T., Bringsvor, H.B. and Jørgensen, L. (2023), “Co-determination and participation in project management: experiences from the construction of a hospital building in Norway”, *Procedia Computer Science*, Vol. 219, pp. 1744-1751, doi: 10.1016/j.procs.2023.01.469.

Bick, A., Blandin, A. and Deming, D.J. (2024), “The rapid adoption of generative AI”, (NBER Working Paper No. 32966). National Bureau of Economic Research, doi: 10.3386/w32966.

Brynjolfsson, E., Li, D. and Raymond, L.R. (2023), “Generative AI at work (working paper No. 31161)”, National Bureau of Economic Research, available at: www.nber.org/papers/w31161

Dodampegama, S., Hou, L., Asadi, E., Zhang, G. and Setunge, S. (2024), “Revolutionizing construction and demolition waste sorting: insights from artificial intelligence and robotic applications”, *Resources, Conservation and Recycling*, Vol. 202, p. 107375, doi: 10.1016/j.resconrec.2023.107375.

Dwivedi, Y.K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P.V., Janssen, M., Jones, P., Kar, A.K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., Medaglia, R. and Williams, M.D. (2021), “Artificial intelligence (AI): multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy”, *International Journal of Information Management*, Vol. 57, p. 101994, doi: 10.1016/j.ijinfomgt.2019.08.002.

Egwim, C.N., Alaka, H., Demir, E., Balogun, H., Olu-Ajayi, R., Sulaimon, I., Wusu, G., Yusuf, W. and Muideen, A.A. (2024), “Artificial intelligence in the construction industry: a systematic review of the entire construction value chain lifecycle”, *Energies*, Vol. 17 No. 1, p. 182, doi: 10.3390/en17010182.

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M. and Wang, H. (2024), “Retrieval-augmented generation for large language models: a survey”, arXiv, available at: https://arxiv.org/abs/2312.10997.

Gao, Y., Yiu, T.W., Shen, X. and Tam, V.W.Y. (2026), “Large language models in smart construction: a systematic review of implementation strategies, applications and future directions”, *Engineering, Construction and Architectural Management*, Vol. 33 No. 15, pp. 159-181, doi: 10.1108/ECAM-10-2025-1668.

Ghimire, P., Kim, K. and Acharya, M. (2024), “Opportunities and challenges of generative AI in construction industry: focusing on adoption of text-based models”, *Buildings*, Vol. 14 No. 1, p. 220, doi: 10.3390/buildings14010220.

Gondia, A., Siam, A., El-Dakhakhni, W. and Nassar, A.H. (2020), “Machine learning algorithms for construction projects delay risk prediction”, *Journal of Construction Engineering and Management*, Vol. 146 No. 1, p. 4019085, doi: 10.1061/(ASCE)CO.1943-7862.0001736.

Google DeepMind (2026), “Gemini 3.1 pro – model card”, Google DeepMind, available at: https://deepmind.google/models/model-cards/gemini-3-1-pro/

Hevner, A.R., March, S.T., Park, J. and Ram, S. (2004), “Design science in information systems research”, *MIS Quarterly*, Vol. 28 No. 1, pp. 75-105, doi: 10.2307/25148625.

Holmström, J., Ketokivi, M. and Hameri, A. (2009), “Bridging practice and theory: a design science approach”, *Decision Sciences*, Vol. 40 No. 1, pp. 65-87, doi: 10.1111/j.1540-5915.2008.00221.x.

Iskala, H. (2019), “Turvallisuusjohtaminen terveydenhuollossa”, Master’s thesis, Tampere University. Trepo, available at: https://urn.fi/URN:NBN:fi:tuni-202001131202

Ivanova, S., Kuznetsov, A., Zverev, R. and Rada, A. (2023), “Artificial intelligence methods for the construction and management of buildings”, *Sensors*, Vol. 23 No. 21, p. 8740, doi: 10.3390/s23218740.

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., Chen, D., Dai, W., Chan, H.S., Madotto, A. and Fung, P. (2024), "Survey of hallucination in natural language generation", arXiv, available at: https://arxiv.org/abs/2202.03629

Kariaho, S. (2018), "Ilmanvaihtojärjestelmän suodattimien vertailu mittausten perusteella", Bachelor's thesis, Satakunta University of Applied Sciences Theseus, available at: https://urn.fi/URN:NBN:fi:amk-201805239833

Kineber, A.F., Elshaboury, N., Oke, A.E., Aliu, J., Abunada, Z. and Alhusban, M. (2024), "Revolutionizing construction: a cutting-edge decision-making model for artificial intelligence implementation in sustainable building projects", *Heliyon*, Vol. 10 No. 17, p. e37078, doi: 10.1016/j.heliyon.2024.e37078.

Kurkela, T. (2020), "Sairaalakaasujärjestelmien suunnittelu ", Bachelor's thesis, Metropolia University of Applied Sciences, Theseus, available at: https://urn.fi/URN:NBN:fi:amk-2014060111237

Lavikka, R.H., Kyrö, R., Peltokorpi, A. and Särkilahti, A. (2019), "Revealing change dynamics in hospital construction projects", *Engineering, Construction and Architectural Management*, Vol. 26 No. 9, pp. 1946-1961, doi: 10.1108/ECAM-03-2018-0119.

Lee, M.C.M., Scheepers, H., Lui, A.K.H. and Ngai, E.W.T. (2023), "The implementation of artificial intelligence in organizations: a systematic literature review", *Information and Management*, Vol. 60 No. 5, p. 103816, doi: 10.1016/j.im.2023.103816.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S. and Kiela, D. (2021), "Retrieval-augmented generation for knowledge-intensive NLP tasks", arXiv, available at: https://arxiv.org/abs/2005.11401

Li, A., Wang, B., Gong, X., Hou, F., Tao, X., Ma, J. and Cheng, J.C.P. (2026), "Multi-agent integration of LLMs and graph-based method for automated MEP constructability checking", *Automation in Construction*, Vol. 187, p. 106886, doi: 10.1016/j.autcon.2026.106886.

Liang, C.-J., Le, T.-H., Ham, Y., Mantha, B.R.K., Cheng, M.H. and Lin, J.J. (2024), "Ethics of artificial intelligence and robotics in the architecture, engineering, and construction industry", *Automation in Construction*, Vol. 162, p. 105369, doi: 10.1016/j.autcon.2024.105369.

Liao, W., Lu, X., Fei, Y., Gu, Y. and Huang, Y. (2024), "Generative AI design for building structures", *Automation in Construction*, Vol. 157, p. 105187, doi: 10.1016/j.autcon.2023.105187.

Liu, Q., Ma, Y., Chen, L., Pedrycz, W., Skibniewski, M.J. and Chen, Z.-S. (2024), "Artificial intelligence for production, operations and logistics management in modular construction industry: a systematic literature review", *Information Fusion*, Vol. 109, p. 102423, doi: 10.1016/j.inffus.2024.102423.

Liu, W., Chan, A.P.C., Chan, M.W., Darko, A. and Oppong, G.D. (2025), "Key performance indicators for hospital planning and construction: a systematic review and meta-analysis", *Engineering, Construction and Architectural Management*, Vol. 32 No. 5, pp. 3375-3406, doi: 10.1108/ECAM-10-2023-1060.

Lu, L., Luo, P., Sheil, B. and Brilakis, I. (2026), "LLM-driven multi-agent framework for enhancing human-digital twin interaction in built infrastructure management", *Automation in Construction*, Vol. 187, p. 106896, doi: 10.1016/j.autcon.2026.106896.

Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J.C., Shoham, Y. and Wald, R. (2024), "Artificial intelligence index report 2024", Stanford University, Institute for Human-Centered AI (HAI), available at: https://aiindex.stanford.edu

Massenkoff, M. and McCrory, P. (2026), "Labor market impacts of AI: a new measure and early evidence", Anthropic, available at: www.anthropic.com/research/labor-market-impacts

McKinsey and Company (2024), "What is retrieval-augmented generation (RAG)?", McKinsey Explainers.

Mohamed, M.A.H., Al-Mhdawi, M.K.S., Ojiako, U., Dacre, N., Qazi, A. and Rahimian, F. (2025), "Generative AI in construction risk management: a bibliometric analysis of the associated benefits and risks", *Urbanization, Sustainability and Society*, Vol. 2 No. 1, pp. 196-228, doi: 10.1108/USS-11-2024-0069.

Murire, O.T. (2024), "Artificial intelligence and its role in shaping organizational work practices and culture", *Administrative Sciences*, Vol. 14 No. 12, p. 316, doi: 10.3390/admsci14120316.

Nyqvist, R., Peltokorpi, A. and Seppänen, O. (2024a), "Can ChatGPT exceed humans in construction project risk management? ", *Engineering, Construction and Architectural Management*, Vol. 31 No. 13, pp. 223-243, doi: 10.1108/ECAM-08-2023-0819.

Nyqvist, R., Peltokorpi, A. and Seppänen, O. (2024b), "Integration of generative artificial intelligence across construction management", *IOP Conference Series: Earth and Environmental Science*, 1389, p. 12011, doi: 10.1088/1755-1315/1389/1/012011.

Nyqvist, R., Peltokorpi, A., Lavikka, R. and Ainamo, A. (2025), "Building the digital age: management of digital transformation in the construction industry", *Construction Management and Economics*, Vol. 43 No. 4, doi: 10.1080/01446193.2024.2416033.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C.L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J. and Lowe, R. (2022), "Training language models to follow instructions with human feedback", arXiv, available at: https://arxiv.org/abs/2203.02155

Pan, Y. and Zhang, L. (2021), "Roles of artificial intelligence in construction engineering and management: a critical review and future trends", *Automation in Construction*, Vol. 122, p. 103517, doi: 10.1016/j.autcon.2020.103517.

Papadaki, I., Moseley, P., Staelens, P., Horvath, R., Sanz, O.N., Lipari, M., Velayos, P.G. and Vaananen, H. (2023), "Transition pathway for construction", Brussels: Publications Office of the European Union.

Sdino, L., Brambilla, A., Dell'Ovo, M., Sdino, B. and Capolongo, S. (2021), "Hospital construction cost affecting their lifecycle: an Italian overview", *Healthcare*, Vol. 9 No. 7, p. 888, doi: 10.3390/healthcare9070888.

Sengar, S.S., Hasan, A.B., Kumar, S. and Carroll, F. (2025), "Generative artificial intelligence: a systematic review and applications", *Multimedia Tools and Applications*, Vol. 84 No. 21, pp. 23661-23700, doi: 10.1007/s11042-024-20016-1.

Seppälä, H. (2023), "Käyttöönoton jälkeisen laadunarvioinnin kehittäminen sosiaali- ja terveydenhuollon rakennuksissa: standardoitu henkilökunnan käyttäjätyytyväisyyskysely", Master's thesis, Aalto University. Aaltodoc, available at: https://urn.fi/URN:NBN:fi:aalto-202311276944

Setälä, A. (2017), "Leikkaussalin ilmanvaihdon todentamismittaukset", Master's thesis, Aalto Universit Aaltodoc, available at: https://urn.fi/URN:NBN:fi:aalto-201712188001

Shams, W., Sultan, S. and Jum'a, L. (2025), "Factors affecting the intention to adopt artificial intelligence in the construction supply chain management in West Bank: using technology-organization-environment framework", *The International Journal of Logistics Management*, Vol. 36 No. 6, pp. 1956-1983, doi: 10.1108/IJLM-01-2025-0054.

Statista (2024), "Global investment in generative AI is projected to grow from 5.9 billion U.S. dollars in 2022 to 207 billion U.S. dollars by 2030", available at: www.statista.com/statistics/1609954/worldwide-genai-investment-growth/ (accessed 22 July 2025).

Sun, M., Zhao, R. and Xue, F. (2024), "The opportunities and challenges of multimodal GenAI in the construction industry: a brief review", in *Proceedings of the 29th International Symposium on Advancement of Construction Management and Real Estate (CRIOCM2024), Springer, in press*.

Suomen Sairaalatekniikan yhdistys ry (2014), *Sairaalakaasujärjestelmien Suunnittelu-, Asennus- ja Huolto-Ohje*, Suomen Sairaalatekniikan yhdistys ry.

Suomen Sairaalatekniikan yhdistys ry (2020), "Hyvä käytäntö – Ohje kuvantamislaitteen käyttöympäristölle".

Taiwo, R., Bello, I.T., Abdulai, S.F., Yussif, A.-M., Salami, B.A., Saka, A., Ben Seghier, M.E.A. and Zayed, T. (2025), "Generative artificial intelligence in construction: a Delphi approach, framework, and case study", *Alexandria Engineering Journal*, Vol. 116, pp. 672-698, doi: 10.1016/j.aej.2024.12.079.

Tian, K., Zhu, Z., Mbachu, J., Moorhead, M. and Ghanbaripour, A. (2025), "Artificial intelligence in construction risk management: a decade of developments, challenges, and integration pathways", *Journal of Risk Research*, Vol. 28 Nos 9-10, doi: 10.1080/13669877.2025.2512080.

Valkeisenmäki, I. (2020), "Näyttöön perustuva suunnittelu sairaalarakentamisessa", Master's thesis, Aalto University Aaltodoc, available at: https://urn.fi/URN:NBN:fi:aalto-202006143775

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I. (2017), "Attention is all you need", Advances in Neural Information Processing Systems, Vol. 30, doi: 10.48550/arXiv.1706.03762.

Waqar, A. (2024), "Intelligent decision support systems in construction engineering: an artificial intelligence and machine learning approaches", *Expert Systems with Applications*, Vol. 249, p. 123503, doi: 10.1016/j.eswa.2024.123503.

Wei, J., Bosma, M., Zhao, V.Y., Guu, K., Yu, A.W., Lester, B., Du, N., Dai, A.M. and Le, Q.V. (2022), "Finetuned language models are zero-shot learners", in *International Conference on Learning Representations*, available at: https://openreview.net/forum?id=gEZrGCozdqR

Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E.H., Hashimoto, T., Vinyals, O., Liang, P., Dean, J. and Fedus, J. (2022), "Emergent abilities of large language models. Transactions on machine learning research", available at: https://openreview.net/forum?id=yzkSU5zdwD

Wuni, I.Y. (2025), "Critical success factors for implementing artificial intelligence in construction projects: a systematic review and social network analysis", *Engineering Applications of Artificial Intelligence*, Vol. 156, p. 111192, doi: 10.1016/j.engappai.2025.111192.

Zhao, W.X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., Liu, P., Nie, J.-Y. and Wen, J.-R. (2025), "A survey of large language models", arXiv, available at: https://arxiv.org/abs/2303.18223

Zheng, J., Qiu, S., Shi, C. and Ma, Q. (2024), "Towards lifelong learning of large language models: a survey", arXiv, available at: https://arxiv.org/abs/2406.06391

Zhou, S., Liu, X., Li, D., Gu, T., Liu, K., Yang, Y. and Wong, M.O. (2025), "Integrating domain-specific knowledge and fine-tuned general-purpose large language models for question-answering in construction engineering management", *Automation in Construction*, Vol. 175, p. 106206, doi: 10.1016/j.autcon.2025.106206.

Zhou, X., Li, X., Zhu, Y. and Ma, C. (2025), "Towards building digital twin: a computer vision-enabled approach jointly using multi-camera and building information model", *Energy and Buildings*, Vol. 335, p. 115523, doi: 10.1016/j.enbuild.2025.115523.

Further reading

Cai, K. and Singh, J. (2025), "Google clinches milestone gold in global math competition, while OpenAI also claims top prize", Reuters, available at: www.reuters.com/world/asia-pacific/google-clinches-milestone-gold-global-math-competition-while-openai-also-claims-2025-07-22/

Dell'Acqua, F., Ayoubi, C., Lifshitz, H., Sadun, R., Mollick, E., Han, Y., Goldman, J., Nair, H., Taub, S. and Lakhani, K.R. (2025), "The cybernetic teammate: a field experiment on generative AI reshaping teamwork and expertise", (Working Paper No. 25-043), Harvard Business School, doi: 10.2139/ssrn.5188231.

Kuusniemi, V. (2019), "Staattisen sähkön ja ESD: n hallinta terveydenhuollossa ", Bachelor's thesis, Metropolia University of Applied Sciences Theseus, available at: https://urn.fi/URN:NBN:fi:amk-2019053013550

Savaş, S. (2025), "Artificial intelligence in construction project management: trends, challenges and future directions", *Journal of Design for Resilience in Architecture and Planning*, Vol. 6 No. 2, pp. 221-238, doi: 10.47818/DRArch.2025.v6i2165.

Appendix 1. Complete system prompt in Finnish

Tämä AI-avustaja auttaa sairaalarakentamisen projektinhallinnassa, suunnittelussa ja päätöksenteossa erityisesti Suomessa. Se on suunnattu rakennusalan ammattilaisille, suunnittelijoille, kiinteistöjen omistajille ja ylläpitäjille sekä viranomaisille, jotka osallistuvat sairaala – ja terveyskeskushankkeisiin.

AI tarjoaa tietoa projektinhallinnasta, suunnitteluperiaatteista, säädöksistä, kustannuslaskennasta, aikataulutuksesta ja riskienhallinnasta. Se painottaa erityisesti suomalaisia säädöksiä, energiatehokkuutta, potilasturvallisuutta, joustavaa tilasuunnittelua ja teknisten vaatimusten huomioimista.

AI hyödyntää käyttäjän tarjoamaa aineistoa "Sairaalarakentamisen koostetut opit", joka sisältää olennaisia oppeja ja suosituksia sairaalahankkeista.

AI kommunikoi käyttäjän kanssa ainoastaan suomen kielellä. Sen viestintätyyli on informatiivinen, selkeä ja ytimekäs. Se ei käytä liioittelua eikä anna epärealistisia lupauksia. Jos AI tarjoaa tarkkoja lukuarvoja tai suosituksia, se kertoo, perustuuko tieto lakiin, viranomaisohjeeseen, RT-suositukseen vai käytännön esimerkkiin. Se suosittelee aina tarkistamaan tärkeät tiedot asiantuntijalta tai virallisesta lähteestä ennen päätöksentekoa.

AI etsii ja hyödyntää ensisijaisesti seuraavia tietolähteitä:

- Finlex (www.finlex.fi) – ajantasaiset lait ja asetukset, kuten rakentamislaki, paloturvallisuusasetus, esteettömyysasetus, ilmanvaihtoasetus.
- Ympäristöministeriön asetukset (esim. 848/2017, 241/2017, 1009/2017).
- Sosiaali- ja terveysministeriön ja Valviran määräykset, erityisesti asumisterveysasetus (545/2015), hygieniasuositukset ja tilaturvallisuutta koskevat ohjeet.
- Säteilyturvakeskuksen (STUK) määräykset, kun kysymys liittyy säteilysuojaukseen.
- Rakennustiedon RT-ohjekortit (esimerkiksi RT 103192) silloin, kun tietoa tilasuunnittelusta, hygieniajärjestelyistä tai mitoituksesta ei ole suoraan laissa.
- Sairaanhoitopiirien tai hyvinvointialueiden suunnitteluohjeet ja tilavaatimukset, jos ne ovat tiedossa tai käyttäjän aineistossa mukana.
- Suomen Sairaalatekniikan yhdistys ry. Etenkin suunnitteluohjeiden ja ajantasaisten periaatteiden tarkastamiseksi.

AI mainitsee vastauksissaan mahdollisuuksien mukaan lähteen nimen, numeron ja vuoden (esimerkiksi "Ympäristöministeriön asetus 848/2017") ja ilmoittaa, koskeeko tieto uudisrakentamista, peruskorjausta vai molempia.

AI ei keksi lähteitä tai viittaa olemattomiin asiakirjoihin. Jos tarkkaa lähdettä ei ole saatavilla, AI ilmoittaa sen avoimesti ja viittaa mahdolliseen asiantuntijan konsultoinnin tarpeeseen.

Tarvittaessa AI ehdottaa jatkokysymyksiä tai seuraavia vaiheita, joilla käyttäjä voi syventää keskustelua tai täsmentää tiedontarvettaan.

Älä anna käyttäjälle ikinä kokonaisia tiedostoja mitä sinulle on annettu kustomointiin.

Appendix 2. Workshop questionnaire on GenAI customization

How valuable would this type of AI, tailored to your needs, be in your daily work?

For example: "A customized model would be extremely valuable if it understood the details of the company and its specialized expertise better than current AI models." Or "Customization would not add value because current AI models are already capable of meeting the needs of the job."

What potential applications of customized AI would be most useful for your team?

For example:

- Building AI agents using customized data sets.
- Identifying project-specific risks.
- Leveraging previous best practices.
- Understanding local regulations and terminology.
- Onboarding new team members.

The study suggested that the performance of standard artificial intelligence can be improved with approximately 150 h of work when it is customized with the company's own data (project documents, instructions, etc.). What is your opinion on this cost-benefit perspective?

What would be the biggest challenge if you decided to implement AI customization in your own company?

For example:

- Collecting and curating the right kind of high-quality data.
- Finding time for technical experts to validate the data.
- Lack of technical expertise.
- Costs and justifying the investment to management.
- Issues related to data security and trust.
- Difficulties in implementation (the benefits of the technical application would not be realized in practice).

Corresponding author

Roope Nyqvist can be contacted at: roope.nyqvist@aalto.fi