DEMOCRATIZING DEEP EXPERTISE: A FRAMEWORK FOR EXTRACTING AND CODIFYING TACIT KNOWLEDGE USING LARGE LANGUAGE MODELS

Irshad Abdulla

Beedie School of Business, Simon Fraser University, Vancouver, Canada

ABSTRACT

The proliferation of large language models (LLMs) opens new possibilities for capturing and scaling human expertise. Yet most applications focus on synthesizing large, documented corpora (papers, reports, manuals, articles, logs). This paper presents a novel approach for extracting and codifying undocumented, deep expertise—the "golden nuggets" of human insight—and making it accessible via a Retrieval-Augmented Generation (RAG) system. Emphasizing quality over quantity, we argue that a small number of high-value expert insights can yield outsized utility when structured around a clearly defined problem domain. Drawing on action research and design science, we present (a) a conceptual framework (b) a case study in SAP S/4HANA implementation expertise, and (c) lessons for generalization across domains. We conclude that combining SME insight with LLMs can democratize scarce knowledge and generate significant value for organizations.

KEYWORDS

Knowledge Management, Commoditization of Expertise, Knowledge Engineering, Artificial Intelligence, Large Language Models, Tacit Knowledge Extraction

1. INTRODUCTION

Organizations increasingly face a paradox: while domain experts hold deep insights critical for success, much of that knowledge remains *undocumented*—residing tacitly in their heads, formed by years of lived experience. Traditional knowledge management (KM) tools (knowledge bases, intranets, manuals) struggle to capture these insights, especially when expertise is scattered, contextual, or informal. Further, domain experts often prioritize the tasks and deliverables required as part of their "day jobs" and comprehensively documenting their deep insights for purposes of knowledge management falls to the bottom of the priority list.

At the same time, LLMs and AI systems are transforming how we interact with information. However, the majority of AI deployments in knowledge domains rely on extensive, preexisting corpora. Studies like *"Towards an AI-based knowledge assistant for goat farmers"* (Han et al., 2025) and *"Retrieval-Augmented Generation to Generate Knowledge Assets and Develop Actionable Insights"* (James et al., 2025) demonstrate the potential of RAG when large, generally well-structured datasets exist. But these approaches depend on *data abundance*, not *insight scarcity*.

David C. Wyld et al. (Eds): MLNLP, ASOFT, CSITY, NWCOM, SIGPRO, AIFZ, ITCCMA – 2025
pp. 165-180, 2025. CS & IT - CSCP 2025 DOI: 10.5121/csit.2025.152013

The core proposition of this paper is that high-quality, undocumented expert knowledge—"golden nuggets"—can be elicited, structured, and made operational quickly, even in data-poor environments. We emphasize that *quality—not quantity*—should be the guiding principle. Moreover, a clear **problem domain definition** is essential to ensure that the knowledge product is relevant, precise, and useful. For example, a knowledge system could be about general consumer behavior, or it could be narrowly specialized—e.g. buying behavior of children's toy cars during holiday seasons.

This study addresses the following research question:

> **How can Retrieval-Augmented Generation (RAG) be applied to accelerate the codification and accessibility of undocumented, deep expert knowledge?**

To explore this, we report on an action research project in which a prototype "SAP S/4HANA Implementation Consultant" AI advisor was developed. Expert interviews yielded deep insights into the problem domain (i.e. the migration from the legacy SAP ECC ERP technology to the new SAP S/4HANA ERP platform), which were structured using an ontology of IT project challenges and connected to LLMs (OpenAI GPT). Through iterative cycles of development and reflection, we derive a **conceptual framework**, evaluate system properties, and highlight implications for practice and further research.

The contributions are threefold:

1. A methodological framework for eliciting and codifying undocumented expertise into RAG systems.
2. A demonstration that *small, high-value datasets* can achieve useful AI guidance when structured around problem domains.
3. Insights into how combining SME insight with LLMs can democratize specialized knowledge and add value to organizations.

2. Literature Review

2.1. Tacit Knowledge, Externalization, and the SECI Model

Tacit knowledge, by Polanyi's (1966) definition, is "personal, context-based, and thus hard to formalize." Nonaka and Takeuchi (1995) introduced the SECI model, positing that organizations create knowledge through four modes: Socialization, Externalization, Combination, and Internalization. The transition from *tacit to explicit* (Externalization) is often the most challenging. Most KM efforts and tools have traditionally relied on documents, wikis, and databases to capture explicit knowledge. This paper aims to use AI as a mediator of externalization and combination, automating parts of the knowledge conversion cycle.

2.2. Knowledge Commoditization and KM Systems

In KM theory, the challenge has always been not just capturing knowledge but making it usable and scalable (Davenport & Prusak, 1998; Alavi & Leidner, 2001). The notion of *commoditizing expertise* refers to converting domain insight into reusable, queryable assets. AI and LLMs are increasingly seen as enablers of a new generation of knowledge systems (von Krogh, 2023; Del Giudice, 2023) that allow organizations to treat expert insight as a service.

2.3. Retrieval-Augmented Generation for Knowledge Systems

RAG architectures (Lewis et al., 2020) combine a retrieval component (fetching relevant passages or chunks) with a generative model (LLM) that composes responses. RAG has been used in open-domain QA, domain-specific systems, and business intelligence (Gao et al., 2023; Huang et al., 2024; Li et al., 2024). However, usage tends to assume large, indexed corpora. Few studies focus on **injecting small, carefully curated expert knowledge** into RAG systems as the core knowledge base.

2.4. Problem Domain Definition and Scope in KM Design

Effective knowledge systems are always built with a domain in mind (Checkland, 1981; Hevner et al., 2004). Without boundary definitions and domain parameters, KM artifacts become unfocused or overly general, reducing usefulness. This paper stresses that problem domain definition must be the starting point for knowledge codification.

3. CONCEPTUAL FRAMEWORK

Below is the conceptual framework we propose—the **Knowledge Commoditization Loop—** showing the pipeline from SME insight to accessible knowledge product.

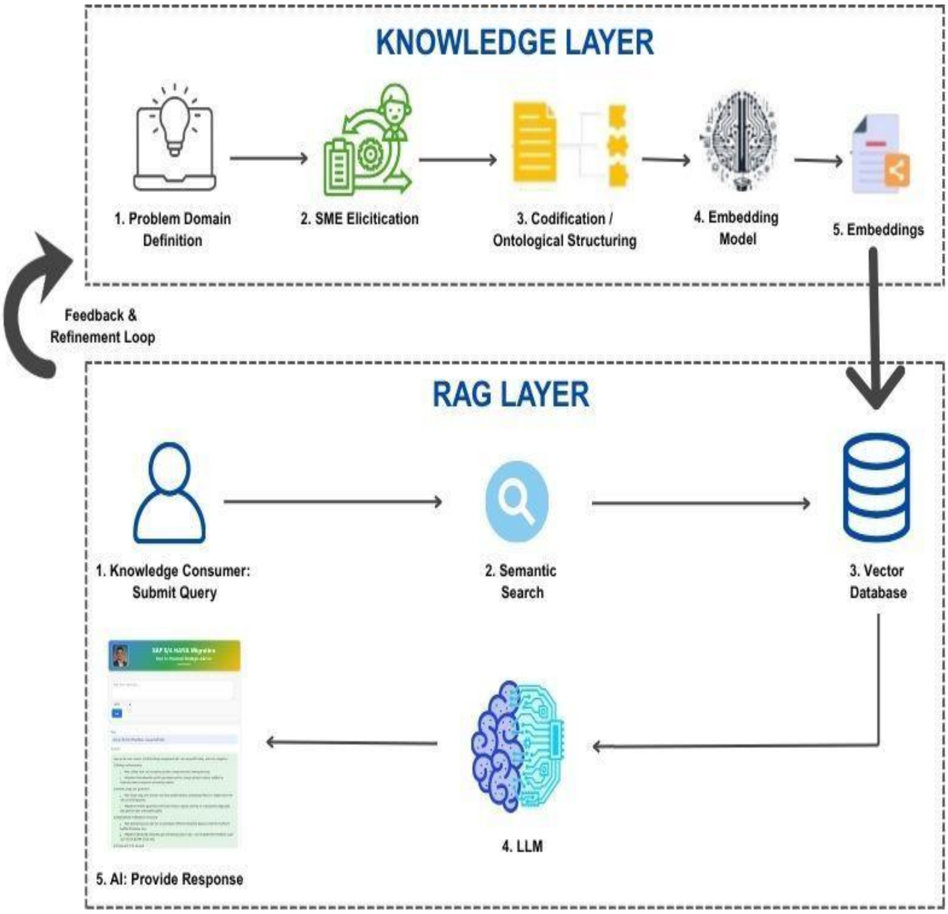

Figure 1. Conceptual Framework of Knowledge Commoditization using RAG
(Adapted from James et al., 2025)

3.1. Key Components

1. **Problem Domain Definition**

    ○ The foundation. Define scope (e.g. domain, context, temporal constraints).
    ○ Determines what kinds of "golden nuggets" are relevant.

2. **SME Elicitation**

    ○ Semi-structured interviews, scenario prompts, reflective questioning.
    ○ Focused on uncovering undocumented, deep expertise often informed by the SME's own lived experience; which is often referred to anecdotally as "been there, done that, got the T-shirt" type of expertise. We are not focused on simply repeating existing documents, or describing well-known theories or methodologies; we are interested in unpacking the nuances, contextual factors and specificities known that provide additional depth to the established body of knowledge in the problem domain.

3. **Codification / Ontological Structuring**

    ○ Map insights into an ontology or taxonomy aligned with the problem domain. This becomes an organizing logic for the "golden nuggets".
    ○ Annotate, chunk, and contextualize each nugget for retrieval.
    ○ Use an embedding model to create embeddings from nuggets and store them in a vector database.

4. **Retrieval + LLM Integration (RAG Layer)**

    ○ Employ a retriever to fetch relevant codified nuggets.
    ○ LLM generates explanations, guidance, or narratives grounded in those nuggets.
    ○ LLM also suggests follow-up questions or guidance, grounded in applicable nuggets. This helps guide the user through a thoughtful and high-value conversation.

5. **Knowledge Consumers**

    ○ Users pose queries; the system returns grounded responses.
    ○ Helps democratize expert insight.

6. **Feedback & Refinement Loop**

    ○ Users or SMEs review outputs, identify gaps or errors.
    ○ Loop back to elicitation or codification to refine and extend knowledge.
    ○ Loop back to problem domain definition to review and adjust parameters of the problem domain; or potentially start a new problem domain.

3.2. Emphasis on Quality over Quantity

This framework pivots away from brute-force scale and pivots away from trying to derive deep insights from the plethora of existing data in the problem domain. Rather than indexing thousands of documents, the methodology prioritizes **carefully elicited, high-value insights** that are logically structured and immediately useful. The system effectiveness depends less on volume and more on the *signal quality and alignment with domain scope*.

4. METHODOLOGICAL APPROACH

This study is framed as **action research / design science**. The researcher is both designer and evaluator, iteratively building and refining a prototype.

4.1. Problem Domain: SAP S/4HANA Implementation Consulting

The case study for this research is situated within the context of enterprise SAP transformations — a domain widely recognized for its complexity, strategic importance, and high failure risk. SAP's legacy ERP system (ECC) is approaching end-of-life (EOL) (support ending in 2027), which prompts a massive global migration effort to the new platform SAP S/4HANA. Many organizations still rely on ECC, and as of late 2024 only about 39% of the estimated 35,000 SAP ECC customers had transitioned to S/4HANA. At current migration rates, forecasts suggest roughly half of ECC installations might remain legacy beyond 2027(CIO, 2025).

This migration wave is fraught with risk. Even with established frameworks and methodologies, many S/4HANA projects suffer budget overruns, delays, and quality shortfalls. A recent industry study (Horváth) indicates that projects typically take 30% longer than planned, and only 8% of organizations complete migrations on time. Budgets are frequently exceeded, and many migrations deliver results that fall short of quality expectations (CIO, 2025). Another survey found that 65% of respondents missed quality targets, and 90% of migrations exceeded original schedule estimates (TTC Global, 2025).

Failure is not due to lack of frameworks alone. ERP and S/4HANA migration literature is rich with critical success factors (CSFs), risk taxonomies, and best practices (e.g. change management, data migration strategy, governance) (OxfordCorp, 2024). Yet these generic guides often fail to anticipate project-specific contingencies that arise due to idiosyncratic legacy architectures, customized modules, data quality issues, organizational culture, strategic decision-making, skills gaps or system integration complexities. Each implementation is unique, and what went well in one environment may not apply in another.

Subject Matter Experts (SMEs) carry valuable experiential knowledge — what works, what fails, subtle signals, "lessons from the trenches" — that often go undocumented. These "golden nuggets" are insights that transcend generic templates and are tuned to domain and context. They may include heuristics about hidden dependencies, temporal sequencing of tasks, integration pitfalls under resource constraints, or mitigations tailored to specific business scenarios.

Compounding this is an acute skills shortage in the SAP space. The demands of S/4HANA migration are far outstripping supply of experienced consultants. Many ECC consultants lack deep S/4HANA experience, and even for those who do, the newer platform features, evolving best practices, and architectural innovations make knowledge fragile. The talent gap has become a barrier to migration itself; for example, nearly one-quarter of surveyed SAP users reported that the shortage of qualified experts delayed projects (Forbes, 2024; ThirdStage, 2024).

AI and LLMs provide a promising mechanism to scale deep expertise by capturing and structuring these golden nuggets, then making them queryable and accessible to a broader audience. In this way, AI becomes not a replacement for experts but a lever to democratize strategic knowledge, especially in domains where human expertise is scarce and high-value.

In sum, the SAP S/4HANA migration domain exemplifies a high-stakes, high-variance transformation where documented knowledge is insufficient and tacit expertise is critical. It is thus an ideal testbed for exploring how to extract, codify, and deliver deep, undocumented insights via RAG systems.

4.2. Elicitation & Interview Protocol

- Expert interviews using guided prompts and scenario-based questioning Encouraged reflection on past projects, anomalies, and edge-case lessons
- Captured insights as raw text, then validated with SMEs

4.3. Ontology / Knowledge Structuring

- Used a pre-existing ontology of IT project challenges (Abdulla & McArthur, 2018)
- Mapped each nugget to ontology nodes, provided context and metadata (project phase, conditions, dependencies)
- Chose chunk structure and chunking size to balance specificity and generality

4.4. System Implementation

- Backend: vector store (e.g. ChromaDB, FAISS / Pinecone) indexing the codified nuggets
- RAG setup: OpenAI GPT models (configurable across GPT-3, GPT-4, GPT-5) used for generation. Python to convert SME inputs into vector database and to orchestrate communication with OpenAI.
- Frontend: query interface (command-line or web-based e.g. REACT and FastAPI) to test responses

4.5. Evaluation & Iteration

- The researcher tested queries, compared outputs across GPT models
- Assessed correctness, coherence, relevance, grounding. Reviewed the RAG outputs against standard ChatGPT outputs to assess depth and value of the RAG outputs.
- Multiple cycles led to prompt tuning, ontology refinements, and nugget additions

4.6. Limitations

- Only one domain was tested
- Evaluation was qualitative and limited to the researcher's judgment
- No large user base or formal metric validation in this phase

5. Case Example & Illustrative Output

This case study assumes that the users of the AI solution are themselves subject matter experts (SMEs) within the defined problem domain. These users seek to access additional deep insights from their peers—insights that may not be widely documented but are valuable for decision-making. The AI solution enables this by retrieving and presenting such insights quickly and intuitively. Given their expertise, end users are able to critically evaluate the AI-generated responses and determine how best to apply them. The solution is therefore designed primarily for experienced practitioners rather than novices.

5.1. User Interface Design

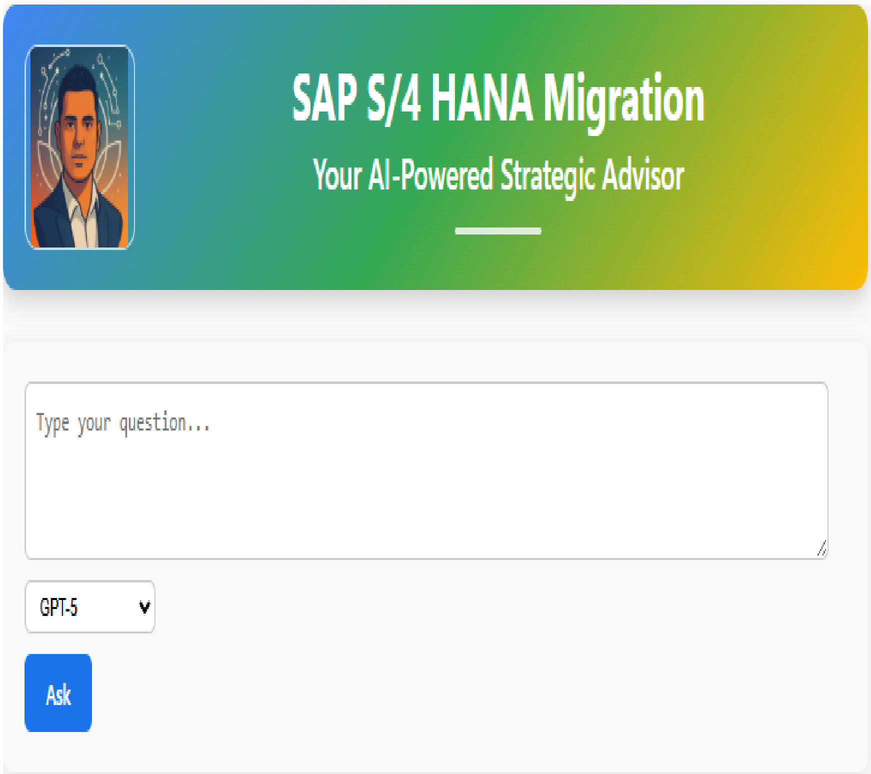

Figure 2. User Interface for RAG-enabled AI solution providing deep expertise

The user interface (Figure 2) follows a minimalist design similar to ChatGPT. This intentional simplicity ensures that users experience minimal learning curve or onboarding friction. Users enter any question related to the problem domain and click the **Ask** button. The AI solution then retrieves relevant "golden nuggets" from the curated knowledge base and constructs a context-enriched prompt, which is passed to the large language model (OpenAI GPT-5). The response is displayed to the user in real time, accompanied by suggested follow-up questions.

Users may (a) ask the suggested follow-up question, (b) tailor it to their specific context, or (c) formulate a new query. Upon completing the session, users can click **End Chat** and optionally download the full conversation transcript for future reference.

5.2. Illustrative Query

To demonstrate the depth and value of insights produced by the RAG-enabled system, we posed the question:

"What are the S/4HANA change management risks?"

We then asked the same question directly to ChatGPT (also powered by GPT-5) and compared the responses. Evaluation was based on the following criteria:

- **Prerequisite:** Each identified risk must be valid, regardless of whether it represents a general or domain-specific concern.

- **Scoring:** One point was awarded for each deep insight—or *golden nugget*—embedded in the explanation of a risk.

5.3. Evaluation

Appendix A presents the detailed comparative analysis. Findings indicate that the RAG-enabled system, grounded in SME-elicited "golden nuggets" and enriched by LLM reasoning, consistently produced contextually grounded, situationally relevant advice that extends beyond generic templated advice.

Each *golden nugget* need not be exhaustively described in the response. As illustrated in

Appendix B, even a concise mention provides valuable "breadcrumbs" that guide the user toward deeper inquiry. In contrast, the standard ChatGPT responses, lacking these embedded nuggets, offered limited guidance and failed to orient the user toward unrecognized areas of potential exploration—illustrating the well-known problem that *"users do not know what they do not know."*

The comparative analysis was repeated 20 times across diverse queries. In all cases, the RAG-enabled solution produced insights that reflected deeper domain expertise and higher perceived value. This supports the argument that Retrieval-Augmented Generation (RAG) can accelerate the codification and accessibility of undocumented, deep expert knowledge, operationalizing the conceptual framework outlined in Section 3.

We acknowledge that identifying what constitutes a *golden nugget* is inherently subjective and, in this study, reflects the researcher's expert judgment. To enhance the credibility and rigor of the findings, future work will involve additional SME reviewers to validate the categorization and interpretation of "golden nuggets" as presented in Appendix A.Future work could also include quantitative usability testsor knowledge retention metricsto measure the effectiveness of LLM-assisted codification.

6. IMPLICATIONS FOR PRACTICE

The proposed approach carries important implications for how organizations capture, curate, and operationalize deep expertise in scalable and ethical ways.

Democratization of Scarce Expertise

By structuring and embedding expert reasoning into retrieval-augmented systems, organizations can make high-value tacit knowledge widely accessible. This reduces dependence on a small group of experts and promotes consistent, evidence-informed decision-making across the organization. Over time, such systems foster *organizational memory*—a collective capability that endures beyond individual tenure. This becomes even more valuable as the workforce ages and access to experts becomes even more difficult.

Speed to Value

Traditional knowledge management or AI programs often require months before benefits materialize. The proposed method enables rapid deployment of "minimum-viable-knowledge" artifacts—turning expert input into usable digital assets within days or weeks. This agility complements iterative and agile implementation practices, providing early business value while allowing progressive refinement. Speed is realized through the simple, yet effective technology infrastructure as illustrated in the prototype solution. Speed is also realized by minimizing amount of expert time required to codify the deep insights and make it usable in the organization as compared to traditional knowledge management methods which require significant expert time and effort.

Scalability and Knowledge Lifecycle Management

Practical adoption demands disciplined lifecycle management. Expertise evolves; therefore, mechanisms to validate, update, and retire nuggets are essential. Governance models should include version control, audit trails, and scheduled review cycles. Large organizations may adopt federated stewardship models, where domain leads own and update their respective knowledge segments, ensuring both scalability and accountability.

Lean and Inclusive Investment

Because the framework relies on structured human cognition rather than vast datasets, it is feasible for small teams and resource-constrained organizations—such as public agencies, regional hospitals, or nonprofit entities. It also democratizes contribution: diverse practitioners can codify their unique insights, broadening representation and enhancing organizational equity.

Ethical and Epistemic Responsibility

Translating tacit expertise into explicit, machine-readable form introduces epistemic risk. Context can be lost, knowledge can be distorted (e.g. through AI hallucinations), nuance flattened, or bias amplified. Use of AI solutions in settings for which the solution was not intended or designed may also introduce risk (E.g. this prototype is not intended for use by novice SAP S/4 HANA users). Implementers must embed safeguards—provenance tracking, transparency logs, peer review, and ethical oversight—to ensure that the democratization of expertise does not compromise its integrity.

Domain Independence and Adaptability

Although demonstrated in an SAP S/4HANA environment, the method generalizes across domains. In healthcare, it could codify clinician heuristics; in energy and utilities, field-engineer judgment; in retail, marketing strategy; in law or policy, interpretive reasoning. The framework thus offers a *meta-methodology* for knowledge democratization applicable wherever deep human expertise is both scarce and consequential.

7. Future Research Direction

Building on these findings, several lines of inquiry can extend both theoretical and practical understanding of RAG-based knowledge systems.

1. **Empirical Validation and Usability Testing**

   Future studies should employ quantitative usability and performance metrics—such as task efficiency, accuracy, user satisfaction, and decision quality—to assess the real-world effectiveness of LLM-assisted codification. Longitudinal designs could evaluate knowledge retention and transfer among users interacting with such systems.

2. **Knowledge Quality and Lifecycle Metrics**

   Research should develop indicators for novelty, coherence, validity, and obsolescence of nuggets. Frameworks for detecting outdated or contradictory knowledge and automated triggers for review will be critical for sustaining accuracy over time.

3. **Hybrid Knowledge Architectures**

   Combining structured nuggets with larger document repositories, relational data, or streaming inputs may yield *layered RAG* architectures that blend human-curated precision with contextual breadth. Comparative analyses can determine optimal balances between curated and emergent knowledge sources.

4. **Automated Elicitation and Codification Aids**

   Advances in NLP and multimodal analytics could support semi-automated extraction of candidate nuggets from interviews, transcripts, or digital traces. Research is needed on how these tools preserve context and mitigate interpretive bias while enhancing efficiency.

5. **Ethical, Epistemic, and Governance Frameworks**

   As AI systems increasingly mediate knowledge, future work must examine how tacit-to-explicit translation affects epistemic integrity. Studies should propose governance models addressing bias, accountability, authorship, and the boundaries between augmentation and automation of expert reasoning. Governance models are also required to ensure that any one AI solution is used only in the setting for which it was designed and tested; and appropriate rigour and due diligence is applied before the AI solution is extended for use in different settings.

6. **Socio-Technical Integration and Change Management**

   Effective implementation depends on human trust, incentives, and workflow alignment. Empirical inquiry into adoption patterns, resistance factors, business process re-engineering and organizational culture will clarify how such systems can be sustainably institutionalized.

7. **Cross-Domain Comparative Evaluation**

   Applying and comparing the framework across sectors—healthcare, energy, legal, retail, education, and public administration—can reveal domain-specific adaptations and extract generalizable design principles for democratizing expertise.

8. CONCLUSIONS

This study proposes a human-centered framework for **democratizing deep expertise** through structured elicitation and retrieval-augmented generation. Instead of pursuing exhaustive data accumulation or end-to-end model training, it foregrounds *curated expert insight*—transforming tacit knowledge into modular, queryable artifacts that large language models can effectively mediate.

The SAP S/4HANA case demonstrates that such systems can be built quickly and deliver immediate value, validating the premise that high-impact organizational knowledge can be captured without massive datasets or specialized technology infrastructure. More broadly, the approach signals a paradigmatic shift: from data-centric AI toward **insight-centric knowledge systems** that prioritize human reasoning, traceability, and domain fidelity.

The implications extend beyond technology. Democratizing expertise entails not only wider access but also **ethical stewardship**—ensuring fidelity, transparency, and accountability in how knowledge is represented and reused. The framework promotes cognitive equity: empowering smaller teams, emerging economies, and under-resourced organizations to leverage AI responsibly while retaining human interpretive control.

Future research should reinforce this foundation through rigorous **empirical validation, lifecycle metrics, and governance models**. Particular attention must be given to preserving contextual meaning, managing knowledge evolution, and defining ethical boundaries for AI-mediated reasoning. By addressing these challenges, scholars and practitioners can advance a new generation of knowledge systems that amplify human insight rather than replace it.

In essence, the path forward lies not in accumulating ever-larger corpora, but in cultivating **structured understanding**—turning individual wisdom into collective capability. Through careful design and stewardship, the democratization of deep expertise can transform how organizations learn, decide, and innovate.

REFERENCES

[1] Abdulla, I., & McArthur, B. (2018). Government-to-government e-government: A case study of challenges facing user adoption in KwaZulu-Natal. Alternation: Interdisciplinary Journal for the Study of the Arts and Humanities in Southern Africa, 25(1), 154–186.

[2] Ahmed, F., Capretz, L. F., Bouktif, S., & Campbell, P. (2015). Soft Skills and Software Development: A Reflection from the Software Industry. arXiv preprint.arXiv

[3] Alavi, M., &Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. MIS Quarterly, 25(1), 107-136.

[4] Checkland, P. (1981). Systems Thinking, Systems Practice. Wiley.

[5] CIO. (2025). SAP users struggle with S/4HANA migration: more than 60% experience deviations in budget, schedule, and quality.CIO

[6] Davenport, T. H., & Prusak, L. (1998). Working Knowledge: How Organizations Manage What They Know. Harvard Business School Press.

[7] Del Giudice, M. (2023). AI and Knowledge Management: The Emerging Role of Generative AI. Journal of Knowledge Management.

[8] Forbes. (2024). The Biggest Tech Talent Gap Can Be Found In The SAP Ecosystem. Forbes

[9] Han, N., Liu, D., & Norton, T. (2025). Towards an AI-based knowledge assistant for goat farmers based on Retrieval-Augmented Generation. *bioRxiv*.

[10] Hevner, A., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. MIS Quarterly, 28(1), 75-105.

[11] Huang, J., et al. (2024). Towards an AI-based knowledge assistant for goat farmers based on Retrieval-Augmented Generation. arXiv preprint.

[12] James, A., Trovati, M., & Bolton, S. (2025). Retrieval-augmented generation to generate knowledge assets and creation of action drivers. *Applied Sciences, 15*(11), 6247.

[13] Kellton. (2025). SAP ECC to S/4HANA Migration Challenges: Preparing for the Future of Intelligent ERP. Blog. Kellton

[14] Li, X., et al. (2024). Retrieval-Augmented Generation to Generate Knowledge Assets and Develop Actionable Insights. Information Systems Frontiers.

[15] Lewis, P., et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. NeurIPS 2020.

[16] Nonaka, I., & Takeuchi, H. (1995). The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation. Oxford University Press.

[17] Oxford Corp. (2024). SAP S/4HANA Migration & Implementations: Overcoming Challenges for Success. Blog / insights.Oxford Global Resources
[18] Polanyi, M. (1966). The Tacit Dimension. Routledge & Kegan Paul.
[19] Reason, P., & Bradbury, H. (2008). The SAGE Handbook of Action Research. Sage.
[20] Resulting IT. (n.d.). Will You Survive The S/4HANA Skills Apocalypse? Skills shortage in SAP ecosystem.resulting-it.com
[21] ThirdStage Consulting. (2024). Addressing the SAP Skill Shortage: Innovations and Strategies.Third Stage Consulting
[22] TTC Global. (2025). Why SAP S/4HANA migrations fail: quality gaps, overruns, schedule delays.TTC Global
[23] von Krogh, G. (2023). Artificial Intelligence and the Democratization of Expertise. MIT Sloan Management Review.

AUTHOR

Irshad Abdulla is an enterprise technology leader with over two decades of experience shaping digital transformation strategies and large-scale architecture practices across healthcare, utilities, consulting, and public sector organizations. He has held senior roles including Chief Technology Architect at a public utility and Director of Digital Transformation at a public health authority. He has also worked in management consulting where he advised clients on maximizing value from enterprise platforms, cloud adoption, and business-technology alignment.

Irshad's research interests lie in better understanding how technology can be applied as a driver of strategic innovation as well as a practical enabler of organizational outcomes. He is a frequent contributor to thought leadership on digital strategy, governance, and enterprise transformation. Irshad holds a PhD in Management Information Systems and Public Administration and teaches at Simon Fraser University, Beedie School of Business.

APPENDIX A

RAG-enabled AI solution: **ChatGPT:**

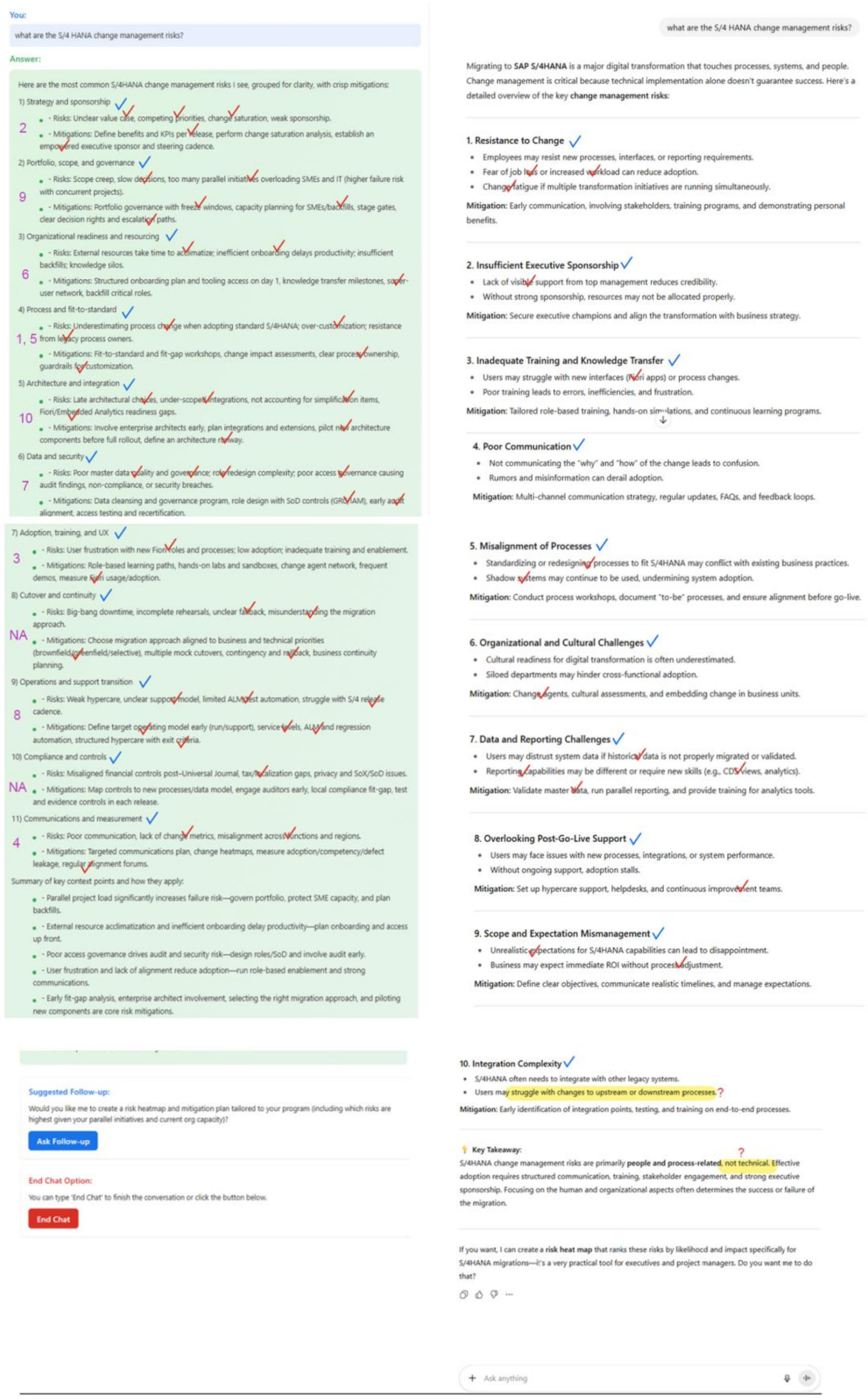

Figure 3. Qualitative comparison of response from RAG-enabled AI solution and ChatGPT

Key:

✓ = **Prerequisite:** Each identified risk must be valid, regardless of whether it represents a general or domain-specific concern

✓ = **Scoring:** One point was awarded for each deep insight—or *golden nugget*—embedded in the explanation of a risk.
Purple Numbering = for ease of reference, the response from the RAG-enabled AI solution is mapped to the closest corresponding response from ChatGPT.
?= AI generated responses that were not clear or questionable.

Analysis:

As indicated by the ✓ annotations, the RAG-enabled AI solution generated **47 golden nuggets** of expert insight, compared to **15** produced by ChatGPT. While these figures might invite a quantitative comparison, such analysis is **not the primary focus** of this study. Moreover, a purely numerical comparison would be misleading, as each golden nugget carries a different level of significance—some insights are far more valuable than others. The perceived importance of a nugget may also vary across users depending on their specific contexts.

Ultimately, the key argument of this research is that **Retrieval-Augmented Generation (RAG)** can accelerate the **codification and accessibility of undocumented, deep expert knowledge**, a process that can be **operationalized** through the conceptual framework presented in **Section 3**

APPENDIX B

The figures below illustrate how the RAG-enabled AI solution generates **"breadcrumbs"**—intermediate insights that guide the user toward **deeper and more meaningful lines of inquiry**.

Figure 3. Breadcrumb from response generated by RAG-enabled AI solution

In the example shown, the original response briefly references *"Fiori/Embedded Analytics readiness gaps"* without further elaboration. Assuming this topic is relevant to the user, it can serve as an entry point for iterative prompting.

As shown below, a concise, one-line follow-up prompt produces an immediate, context-specific response, including a recommended checklist for the user's review. Additional suggested follow-up questions can then be used to explore the topic in greater depth, if desired.

You:

Can you elaborate on Embedded Analytics and what kind of readiness is required?

Answer:

Embedded Analytics in SAP S/4HANA is the set of built-in, real-time analytical capabilities that sit directly on the transactional system (HANA) using CDS views and the SAP Fiori UX. It lets business users analyze live operational data without moving it to a separate data warehouse, through apps such as Analytical List Pages, Overview Pages, Smart Business KPIs, Multidimensional Reports, Query Browser, and Custom Analytical Queries.

What it includes

- - Data model and engine: Virtual Data Model (VDM) built with ABAP CDS views (dimensions, cubes, queries) with annotations for analytics, currency conversion, and semantics. Authorization via CDS DCL and standard S/4 authorization objects.
- - User experience: Fiori analytical apps (ALP, OVP, KPI tiles, charts, tables) with drill-down, filters, and variants. Power users can create queries and KPIs without coding.
- - Tooling: View/Query Browser, Custom Analytical Queries, Manage KPIs and Reports (Smart Business), Analysis for Office integration, and optional SAP Analytics Cloud for advanced dashboards and planning.
- - Extensibility: Add fields to CDS, extend views, build custom analytical queries, or develop new CDS cube/query views in ADT (Eclipse) with transports.

When to use it vs alternatives

- - Use Embedded Analytics for: near real-time operational insights, role-based KPIs, embedded drill-downs in process, light to medium data volumes, limited historical snapshots, and where latency must be minimal.
- - Consider complementing with SAP BW/4HANA or SAP Datasphere + SAP Analytics Cloud for: cross-application or enterprise-wide analytics, historical snapshots, complex transformations, very large data volumes, advanced planning, predictive, or when isolating analytics workload from the transactional system is required.
- - Many clients adopt a hybrid model: embedded for operational analytics inside S/4; enterprise analytics on BW/4HANA/Datasphere + SAC.

Readiness checklist

Business readiness

- - Define "why analytics": prioritize value-driving KPIs and decisions per persona; map existing reports to standard content and identify gaps.
- - Fit-gap: compare legacy reports to standard Fiori analytical apps/CDS content; deprecate low-value custom reports.
- - Ownership and governance: assign KPI owners, data stewards, and a design authority for semantics and drill paths.

Data readiness

- - Master data quality: ensure clean org structures, dates, currencies, units of measure; align valuation and ledger settings (e.g., ACDOCA for Finance).
- - Harmonize dimensions: consistent company codes, plants, customers, materials, profit centers, time-dependent attributes.
- - Currency conversion and time: validate exchange rate types, fiscal calendars, and posting periods used by CDS.
- - Data volume/aging: plan for archiving and data aging to keep operational queries performant.

Technical readiness

- - S/4HANA release and content: confirm your release supports the Fiori analytical apps you plan to use; activate delivered content and roles.
- - Fiori and gateway: deploy Fiori front-end (embedded or hub), enable analytical OData services, spaces/pages, and cache settings.
- - CDS activation and extensibility: ensure standard analytical CDS views are active; set up ADT for CDS development; define naming and transport conventions.
- - Performance and sizing: size HANA for mixed OLTP/OLAP; use CDS @Analytics.dataCategory (#CUBE/#DIMENSION) and @Analytics.query properly; index and restrict default result sets with mandatory filters.
- - Optional components: plan for Analysis for Office or SAP Analytics Cloud (standalone or embedded edition in S/4HANA Cloud) as needed.
- - Embedded BW: keep for delivered content dependencies only; avoid using it as an enterprise DWH.

Security and authorizations

- - Role design: map analytical apps to PFCG roles, Fiori catalogs/spaces; apply least-privilege access.
- - Data control: implement CDS DCL-based row-level security aligned with org structures (e.g., company code, plant).
- - SSO and identities: SAML/IdP integration for Fiori and SAC; manage authorizations consistently across S/4 and any external analytics.

Figure 4. Prompt and response after following the breadcrumb generated by RAG-enabled AI solution