IGNITING KNOWLEDGE MANAGEMENT FOR ASSISTANCE SYSTEMS IN MAINTENANCE: A METHOD FOR KNOWLEDGE GATHERING

Completed Research Paper

Alexander Skolik, Paderborn University, Paderborn, Germany, alexander.skolik@upb.de
Philipp zur Heiden, Paderborn University, Paderborn, Germany, philipp.zur.heiden@upb.de
Tamino Donner, Paderborn University, Paderborn, Germany, tamino.donner@upb.de
Jennifer Priefer, Paderborn University, Paderborn, Germany, jennifer.priefer@upb.de

Abstract

Making tacit knowledge usable for companies has been a challenge. With the latest advancements in generative artificial intelligence (GenAI), developing and using assistance systems has become easier than ever. However, to create assistance systems, an initial base of expert knowledge is required. To establish a method for gathering domain-specific knowledge and especially tacit knowledge, location information, and information systems data, we conducted a design science research study. The designed "IRIS" method captures this information to create an initial spark that serves as the base for assistance systems in the maintenance domain. We demonstrate our method using the case of maintenance in the energy distribution grid and are able to derive an impressive knowledge collection with only a single workshop. The method allows practitioners to gather a wide range of domain-specific knowledge. Researchers in assistance systems and knowledge management can utilize the method for establishing baselines in knowledge collection from different source systems and extensions into GenAI.

Keywords: Assistance Systems, Knowledge Management, Knowledge Collection, Generative Artificial Intelligence, Large Language Models.

1 Introduction

Generative Artificial Intelligence (GenAI) is becoming increasingly important for enterprise information systems. In contrast to non-contextual models, it is expected that over half of GenAI models used in companies will be tailored to specific industries or functions by 2027, up from just 1% in 2023 (Perri, 2024). One common use of context-tailored GenAI models is to blend interactivity with intelligence to create a user-friendly experience—i.e., an assistance system (Piñeiro-Martín et al., 2023). Utilizing text- or voice-based conversational user interfaces, these assistance systems provide intuitive means of communication, offering guidance on complex tasks and answering questions related to various issues (Maedche et al., 2019). For example, an assistance system in power grid maintenance can guide the maintenance technician step-by-step through the maintenance process. The assistance system provides support in the event of uncertainties and responds to queries regarding various aspects of activities during the maintenance process. The assistance system is even able to point out relevant special features, e.g., giving a note on installed components, additional maintenance steps, and location information. For example, high humidity in the areas or prior experiences of employees can be included.

While such an assistance system promises to improve multiple activities in maintenance drastically, setting it up requires different forms and types of data—e.g., location information of the position of assets like switchgears (if they are spread around in a certain area), master data about the asset itself, and transaction data containing protocols of maintenance activities (zur Heiden et al., 2023).

Additionally, business processes can offer insights into the associated activities (Dumas et al., 2018). User-centric knowledge can thus be used to capture knowledge and structure it along the processes (Di Ciccio et al., 2015). However, a significant part of this data is not explicitly available to companies but rather loosely bound in the company as tacit knowledge (Erden et al., 2008; Philipson & Kjellström, 2020). Tacit knowledge, also known as implicit knowledge, resides in individuals' minds and is acquired through experience and practice (McAdam et al., 2007; Polanyi, 2009). Integrating this data provides a comprehensive understanding of the maintenance processes associated with specific assets, enabling companies to tailor assistance systems to their requirements. While companies typically possess a wealth of data, much of the available data remains unstructured, and the absence of externalized tacit knowledge of experts and practitioners limits the establishment of a comprehensive knowledge collection, especially within maintenance processes (Refaiy & Labib, 2009). In addition, in maintenance domain-specific knowledge is needed (zur Heiden et al., 2022). The need for knowledge collection has already been demonstrated in other domains, such as the Internet of Things (IoT) (Santoro et al., 2018). Further, domain-specific knowledge has been found to be fundamental to the practical usage of assistance systems (X. Liu et al., 2023). Therefore, we aim to answer the research question: “How must a method be designed to establish a domain specific knowledge base for assistance systems in maintenance?”This study aims to develop a systematic approach to define, retrieve, and structure the necessary data, information, and knowledge in knowledge-intensive processes such as maintenance processes. The goal is to create a method for gathering and validating knowledge, creating a collection of data that encompasses both explicit and tacit knowledge, customized for the maintenance context. By structuring the collected knowledge, we transform raw, unorganized information into a format that links data points to specific assets, processes, and environmental conditions. This structuring process involves categorizing knowledge artifacts, associating them with their relevant domain-specific attributes, and ensuring they can be effectively retrieved and interpreted by AI systems. Additionally, domain-specific data like location information is integrated to enrich the knowledge base (Fill, 2024; Fill et al., 2024). The data collected from this method can then be structured in a tabular format, which can then be input as part of the knowledge base into the GenAI, which can then assist experts, e.g., maintenance technicians, in their daily maintenance activities. To achieve our design goal, Design Science Research (DSR) is chosen as the research methodology because of its effectiveness in developing and evaluating innovative artifacts to solve complex organizational problems (Hevner et al., 2004). Following the problem formulation, the objective of a coherent knowledge collection is set.

We develop a novel method termed IRIS (Initialization, Retrieval, Investigation, Structuring), designed to collect and structure both explicit and tacit knowledge, as well as location-specific data. We demonstrate the application of the IRIS method by cooperating with a distribution grid operator and are able to extract 169 knowledge artifacts specific to the maintenance of assets on a local distribution grid.

Although various methods show how to collect and organize knowledge within maintenance processes, as for example outlined by Z. Liu and Lu (2024), Motawa and Almarshad (2013), Potes Ruiz et al. (2014) and Sabri (2023), these approaches often focus solely on explicit knowledge. Therefore, they fail to establish a structured collection of tacit knowledge that can effectively customize GenAI to domain-specific contexts. For example, Z. Liu and Lu (2024) define a method for extracting knowledge from maintenance manuals with subsequent modeling of the dependencies between the document elements, but implicit maintenance knowledge is not taken into account. Additionally, they do not incorporate maintenance domain-specific data beyond manuals. Geodata and location information can yield important implications for the maintenance process, especially in power grids (zur Heiden et al., 2023). These data can be acquired, managed, analyzed, and presented with the help of Geographic Information Systems (GIS). The challenge of collecting comprehensive knowledge in maintenance processes is thus raised by incorporating domain-specific source systems, with enterprise resource systems and employee knowledge, rather than a focused, asset-specific view in the maintenance context. This challenge is particular relevant in energy grids since due to the changes in power generation and a lack of skilled workers tacit knowledge is more valuable than ever (Czako, 2020; Smith et al., 2022).

Our research complements existing research in IS by providing a theoretical understanding of how different sources of knowledge—tacit knowledge, master data, transaction data—can be combined to enable GenAI-driven services. Thus, our method enables further GenAI innovations and use cases, in which existing (pre-trained) AI models are not able to adequately respond to context-intensive queries. Researchers in neighboring domains can adapt our method for their data sources, while practitioners can build and extend their knowledge base, necessary for assistance systems using GenAI (Azevedo et al., 2023). Ultimately, our results enable practitioners to build context-specific GenAI applications incorporating organizational knowledge.

The paper is structured as follows. The second section provides an overview of the related work, detailing assistance systems, knowledge management, and location information. Following, we explain the research methodology used. The fourth section presents the design of our IRIS methodology, which features a demonstration and subsequent evaluation. Following the discussion section, the final section concludes our paper.

2 Background and Related Work

2.1 Knowledge Management

Knowledge Management (KM) encompasses the systematic processes of creating, sharing, utilizing, and overseeing knowledge and information within an organization (Girard & Girard, 2015). Therefore, KM should ensure that the appropriate knowledge is provided to the right people at the right time (O'Dell et al., 1998). From an organizational perspective, human knowledge can be classified into the two main categories of explicit knowledge and implicit knowledge (McAdam et al., 2007; Polanyi, 2009). Explicit knowledge, which can be transmitted between individuals or groups using systematic and formal language (Polanyi, 2009). This type of knowledge is typically easy to articulate, capture, and distribute within an organization (Polanyi, 2009). Implicit knowledge, also known as tacit knowledge, is a form of knowledge that is highly personal and context-specific, making it challenging to formalize, codify, and communicate effectively (Polanyi, 2009). Implicit knowledge often resides in the minds of individuals and is acquired through experience and practice, making it a valuable but elusive asset for organizations (McAdam et al., 2007; Polanyi, 2009). These two types of knowledge are not mutually exclusive or independent of one another. Explicit and implicit knowledge exists on a continuum and often interact with each other (Polanyi, 2009). By recognizing and addressing both forms of knowledge, organizations can better harness their collective intellectual capital and drive both innovation and competitive advantage (Szelągowski, 2021; Werrmann & Krämer, 2018).

A potential way to capture explicit knowledge is the creation of knowledge graphs, for example with a task centric approach (Z. Liu & Lu, 2024). Establishing a knowledge base can tie tacit and explicit knowledge together (Motawa & Almarshad, 2013). Feedback from practitioners can expand this knowledge base (Potes Ruiz et al., 2014). This feedback can be gathered from following specific business processes (Sabri, 2023). However none of these methods allows for a methodical approach to gather tacit and explicit knowledge across multiple systems.

To harness the tacit knowledge possessed by experts and practitioners, it must first be externalized into explicit knowledge (Nonaka, 1994). This externalization can take various forms, for example, storytelling and after-action reports (Seghroucheni et al., 2023). While these methods can capture the previously tacit knowledge of experts, there is lack of research combining these externalization methods and data gathering methods, applied on different systems, to one method that is able to collect data and tacit knowledge for GenAI applications.

2.2 Assistance Systems

An assistance system is a technical system designed to receive and process information from its surrounding environment to provide support for the user's needs (Bannat, 2014). Assistance systems can be classified into two categories: basic and advanced assistance systems. Basic assistance systems are defined by their "low degrees of both intelligence and interaction" (Maedche et al., 2016). These systems rely on contextual information and necessitate manual execution. An example of a basic assistance system is the process guidance system, which utilizes simple feedback and usage data to provide basic awareness but lacks contextual understanding (Maedche et al., 2016; Morana et al., 2015). In contrast, advanced assistance systems can determine whether to adhere to the provided assistance and deliver either context-aware or proactive support (Maedche et al., 2016). These systems adapt their functionalities based on user behavior and need (Maedche et al., 2016).

Integrating assistance systems with Artificial Intelligence (AI) seems highly suitable, especially in the service sector (Link et al., 2020). AI-based assistance systems are therefore characterized as highly versatile support tools that can be seamlessly integrated into existing workflows (Peissner et al., 2019). Especially in the domain of maintenance, AI-based assistance systems can have a significant impact and enhance time efficiency (Shin et al., 2021). For example, GenAI Systems can impact the decision support capabilities of assistance systems by providing accumulated best practices and from previous maintenance tasks and stay up to date on safety guidelines. Furthermore, GenAI can be utilized to improve these systems. GenAI is seen to increase the intelligence needed for advanced assistance systems as they employ AI models capable of using learned patterns to generate new data (Feuerriegel et al., 2024). These GenAI applications can then provide solutions to real-world problems, such as creating text, images, speech, or programming code (Feuerriegel et al., 2024). Increasing the intelligence of assistance systems is thus required to establish advanced assistance systems (Maedche et al., 2016). These systems create new opportunities in human-computer interaction and social computing, as the use of natural language enhances usability and accessibility. Thus, impacting current computer-mediated communication and collaboration (Feuerriegel et al., 2024).

For knowledge-intensive processes, the effective tailoring of GenAI assistance systems to domain- or user-specific assistance systems is dependent on suited knowledge collections (Seufert & Meier, 2023). A structured knowledge collection for the development of assistance systems in maintenance is thus necessary (Azevedo et al., 2023). Establishing such a knowledge base can then offer the recommendations and reasonings needed for establishing a well-defined output (Maedche et al., 2019).

2.3 Location Information and GIS

Location information have proven to be particularly valuable for asset management and condition monitoring, for example in the context of energy grid maintenance (zur Heiden et al., 2022). By integrating location information, such as environmental data of weather conditions and pollution levels, a more proactive approach to maintenance is enabled, allowing for early detection of potential equipment failures and enhancing information relevant to the maintenance of an asset. GIS are systems consisting of "hardware, software, data, people, organizations and institutional arrangements for collecting, storing, analyzing and disseminating information about areas of the earth" (Dueker & Kjerne, 1989, pp. 7–8) and enable utilizing location information by acquiring, managing, analyzing, and presenting these data (Ehlers & Schiewe, 2012). Although GIS have been recognized for their strategic potential since the 1990s (Keenan, 1997; Murphy, 1996), their application in the Information Systems (IS) discipline has gained more attention in recent years, particularly regarding big data, mobile computing, and the Internet of Things (IoT) (Elsahlamy et al., 2021; Priefer, 2023; Zhang & Yu, 2014). Today, GIS are used to manage large-scale spatial data collected through sensors, aerial drones, cameras, or mobile devices, as they improve decision-making and enable innovative applications, e.g., in the context of smart grids (Ashkezari et al., 2018; zur Heiden et al., 2022) and IoT (Cao & Wachowicz, 2019).

3 Research Method

To design a structured method for the systematic identification, retrieval, and organization of essential data, information, and knowledge within maintenance processes, we adopt Design Science Research (DSR) as our guiding research paradigm (Hevner et al., 2004). DSR provides a robust structure for creating and evaluating novel solutions to address complex, real-world problems (Hevner et al., 2004). We apply the DSR methodology by Peffers et al. (2007), which comprises the following six steps: identification and motivation, objectives of the solution, design and development, demonstration, evaluation, and discussion and communication. The six steps of the DSR methodology, as applied in our research, are visualized in Figure 1.

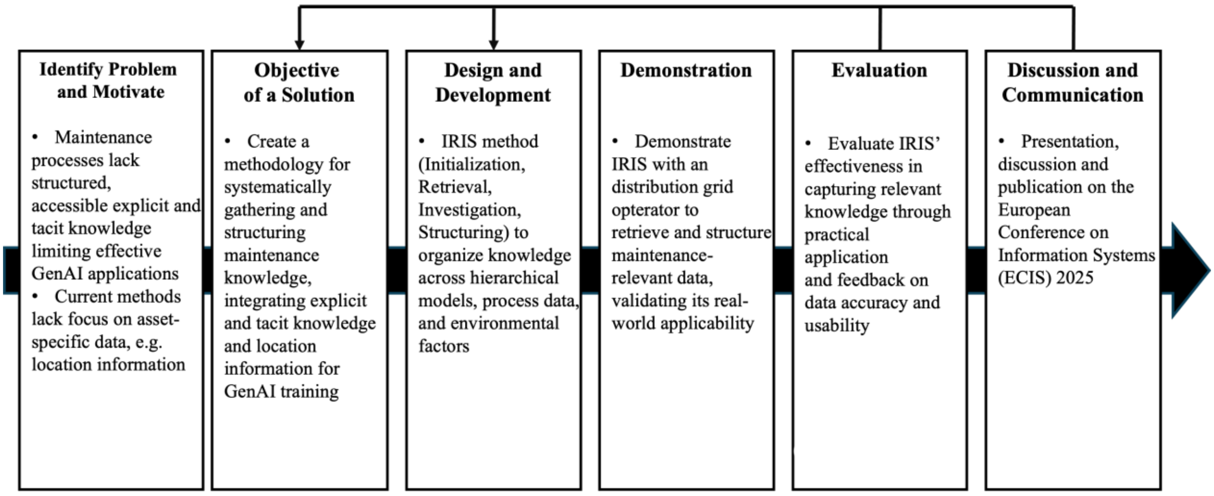

Figure 1. Applied DSR approach adapted from Peffers et al. (2007)

The focus of our research is on maintenance processes for assets on the energy grid. European energy grids are structured across multiple voltage levels to ensure efficient transmission and distribution (BMWK, 2024). Initially, high-output sources like nuclear plants feed electricity into the extra-high voltage (EHV) grid. From there, transformer stations step down EHV to high voltage (HV) for supplying large industrial clients and receiving input from facilities like biogas plants. HV is then reduced to medium voltage (MV) for industrial and service sector use and also accepting contributions from renewable sources like wind and large solar installations. Finally, the low voltage (LV) grid supplies electricity to homes, businesses, and smaller facilities through distribution substations. Key components within these substations, such as switchgears and transformers, are critical for stable energy distribution. Due to aging, pollution, and thermal stress, these components can fail and disrupt power to households, industries, and essential infrastructure. Highlighting the need for robust maintenance practices and set the problem and motivation for our design science research (Hoffmann, 2020; Zickler et al., 2005).

Collecting knowledge that is therefore not only limited to what is present in the system should also include the knowledge and experiences gained by the employees. Already present methods touch on these subjects but are mostly limited by strictly utilizing source systems for their knowledge gathering (Z. Liu & Lu, 2024; Motawa & Almarshad, 2013). Therefore, the objective of the solution is characterized by gathering and structuring systems-, processes- and tacit employee knowledge in maintenance processes.

To do so, the design and development phase extended the theoretical background by adding current literature from the domains of business process management and data science. As a guideline for the creation, the definition of a method is a structured sequence of activities (March & Smith, 1995). The resulting method called "IRIS" thus draws from established concepts and focuses on maintenance specific elements like the utilization of location information. The method was then tested at a distribution grid operator with a total of three online workshops and one workshop with ten employees. Multiple approaches for activities were used during the demonstration and the development of the knowledge base evaluated. The method, alongside the recommendations from this evaluation, is communicated as part of this paper.

4 Knowledge Collection Method IRIS

4.1 Design

The proposed method is titled IRIS as it utilizes these four stages of initialization, retrieval, investigation, and structuring with multiple activities within each stage. The stages are based on various data and process collection methods, especially the BPM life cycle (Dumas et al., 2018), the data science life cycle (Stodden, 2020), the knowledge-based engineering life cycle (Stokes, 2001), and the DASC-PM model (Schulz et al., 2021). These lifecycles all highlight the need for an identification and data acquiring phase leading to the initialization and retrieval phase. Using and formalizing the data are the base for the investigation and structuring phases in the IRIS method. The method is visualized Figure 2 using Business Process Model and Notation (BPMN).

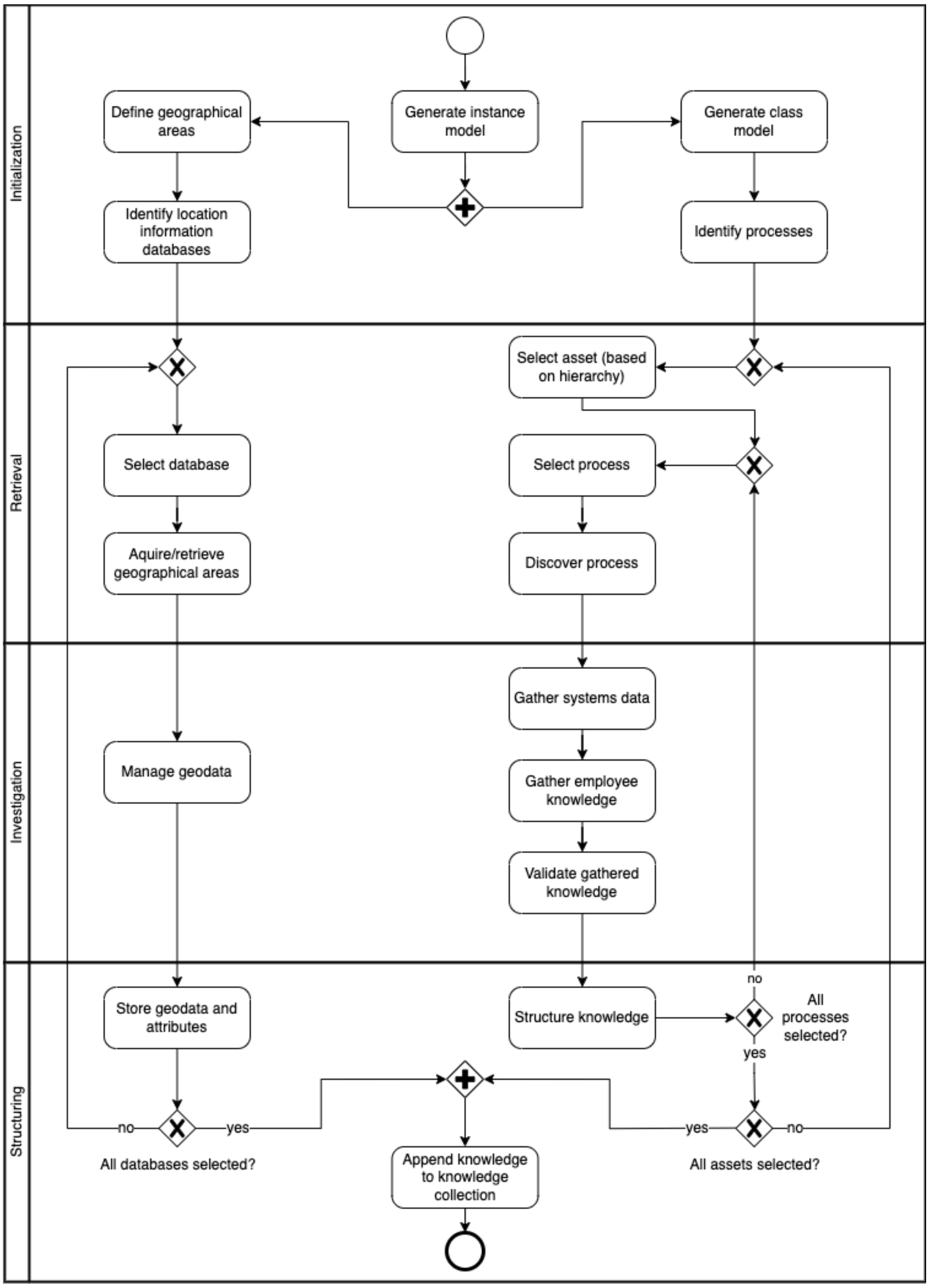

Figure 2. IRIS Method for Knowledge Collection

Within the initialization phase, firstly an instance model of the relevant assets within the company is created. This instance model establishes a base for developing and sourcing knowledge (Wiig, 1998). As part of the instance model, all applicable real-world assets are structured in a hierarchical order, enabling a technical view of each asset and the components it is made of. This allows to gather specific knowledge on a single real-world asset later. From this instance model, two activities deviate.

Once the process has been discovered, the investigation stage commences. Systems data involving the same activities within a process can be gathered from a wide range of systems like enterprise resource planning (ERP) systems or workflow management systems. This data is already structured and can be used as a base for discussions with employees to gather their tacit knowledge. For example, process models can be established and extended to capture the insights from process stakeholders (Dumas et al., 2018). Storytelling enables employees to share insights as stories (Seghroucheni et al., 2023). These stories can then be recorded and establish a collection of tacit knowledge (Gouvêa et al., 2016). After-action reviews aim to gather experiences and learnings by reflecting on the desired process or asset (Seghroucheni et al., 2023). These methods can be done individually or in groups. Once the tacit and explicit knowledge is collected, it has to be validated, i.e., it must be made sure that the gathered knowledge on an individual level matches with the group or company level (García-Fernández, 2015). This can, for example, be done by dividing the employees into multiple groups during the "gather employee knowledge" activity, as this allows to check within and among the groups if the knowledge is correct or potentially harmful (García-Fernández, 2015). Finally, in the structuring stage, the gathered knowledge must be placed in a knowledge collection. As part of the resulting structuring activity, each bundle of data, information, and knowledge has to be paired with an asset and an activity within the analyzed process. Once each bundle within the process and the asset is collected, either further loops on process and asset selection are conducted, or the knowledge is appended to a general knowledge collection. On the left side, the domain-specific data is gathered. In our maintenance context, location information in form of geospatial data are set as the relevant information since they are able to tie location-specific phenomena to an asset. Power grid assets are influenced by various environmental conditions, such as weather, proximity to forests, humidity, animal populations, or pollution. Therfore,the integration of the geographical area is a relevant source of knowledge. While geospatial data are not strictly mandatory for the method, their inclusion significantly enhances the method's effectiveness by enabling detailed analyses of environmental factors and knowledge about common failures, e.g., due to animal intrusion or pollution by pollen count. Therefore—in the initialization phase—relevant geographical areas need to be identified, which serve as a starting point for retrieving the geospatial data from different sources.

Selecting the relevant databases thus serves as a set up activity in the retrieval phase. Since geographical conditions apply to the assets utilizing the instance model, geographical areas can also be derived at this stage for gathering environmental data. Influencing environmental factors are acquired and mapped in form of geospatial data, of which many can already be deprived from internal or external geodatabases or open geospatial data portals (Directive 2007/2/EC, 2007). Integrating these data is crucial for the proposed method, as it enables the incorporation of external environmental influences into the asset management process. GIS enables to acquire, manage, analyze, and present the data (Ehlers & Schiewe, 2012). Defining relevant geographical areas around the assets using GIS to gather the environmental data, allows for a more comprehensive understanding of the conditions impacting the assets. The data is collected and integrated with the help of GIS, which also provides the tools to manage the data in the investigation phase, e.g., transforming coordinates. Within these areas, the proximity to forests and, thus, animal populations can be examined with geographical tools like buffer analysis. In the last phase—structuring—the collected and managed geodata are stored within geodatabases and enriched by attributes, like the type of pollen or times when these pollen disperse and potentially pollute grid assets. Similar to the right side of the method, once all databases are used, the gathered knowledge is appended to a geospatial data collection. Utilizing coordinates of each instance the geospatial data and knowledge collection tables can be related to each other.

4.2 Demonstration

The presented method was demonstrated and tested at a medium-sized distribution grid provider, maintaining its own distribution grid. As the distribution grid provider aims to implement its own GenAI application for maintenance but lacks GenAI application knowledge and does not have an explicit collection of the employees' knowledge, we deem this use case as suitable for demonstrating the use of IRIS.

For the first iteration, maintenance and asset data from the ERP system was utilized to create the instance model. The hierarchies set up in the ERP system could directly be translated into the instance model completing the first activity. The infrastructure of the distribution grid from the whole power grid to the single components like transformers or power panels are hierarchy is visualized in Figure 3.

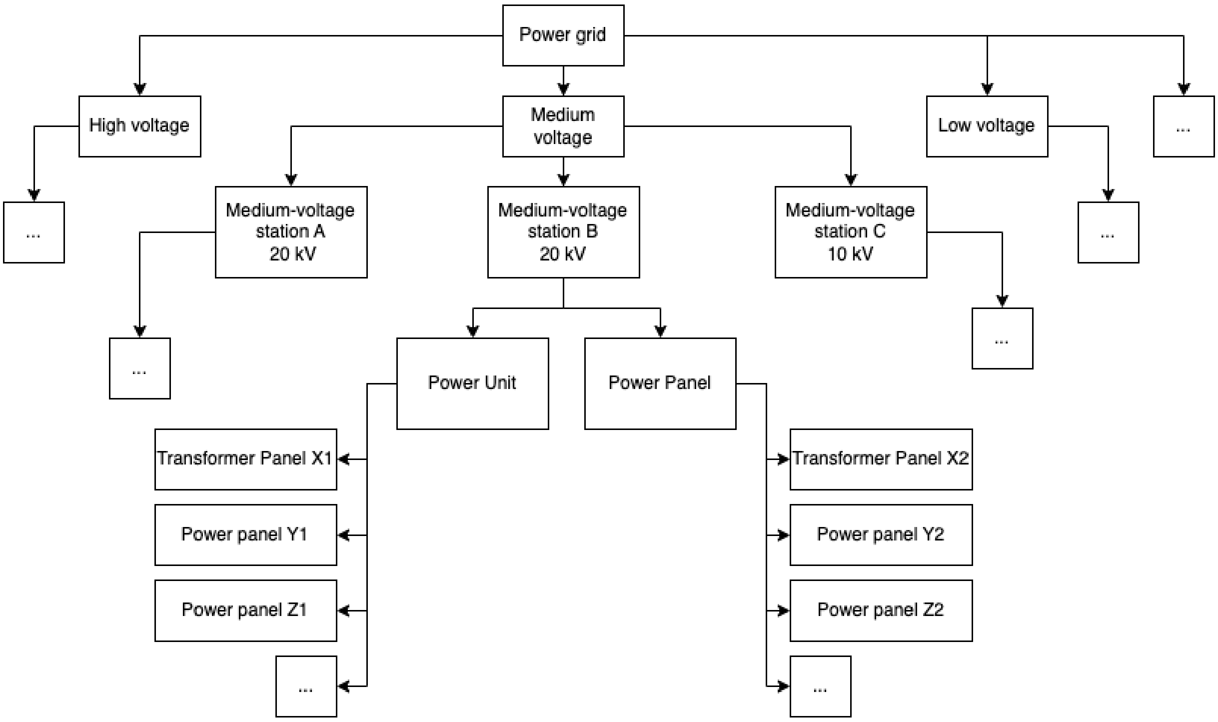

Figure 3. Anonymized Instance Model, highlighting a single Power Station

Based on the instance model, we derived a class model containing different grids and assets of the energy provider. The class model groups instances with similar functions together to create a coherent overview of the different types of assets in the power grid. Figure 4 visualizes our class model from the whole power grid down to different voltage levels and the corresponding assets, like cabinets or power panels in a hierarchical order. Rather than having multiple power stations, a class power station, including the contents of the power stations and their functions, is created. This enables us to gather class-specific knowledge that applies to each instance of the class as well.

The activity "identify processes" was conducted in two ways. First, the potential processes in the ERP system were examined. Second, the business process catalogue of the company was utilized to identify any process set within asset management. Identified processes were, for example, the installation process, malfunction resolution, and the maintenance process. These processes were identified among a total of five core processes and five sub-processes. Together with the class model, we selected one asset and performed the following activities of our IRIS method. Since the maintenance of MV stations is most important to the distribution grid provider, these two were selected as the chosen process and related assets. To discover the process, a process model was present but lacked depth, especially in the crucial parts of the maintenance activities while performing the maintenance. To gather these insights, within the process discovery, three group workshops with maintenance employees and asset managers were conducted to gather additional knowledge on the process activities and associated systems. Thus, it was possible to get an almost step by step view into the maintenance process. This set up the activity of gathering systems data, as it could precisely be determined which information system stores which data. Checklists, maintenance records, and notes were gathered this way and enhanced the data aspect of the knowledge collection.

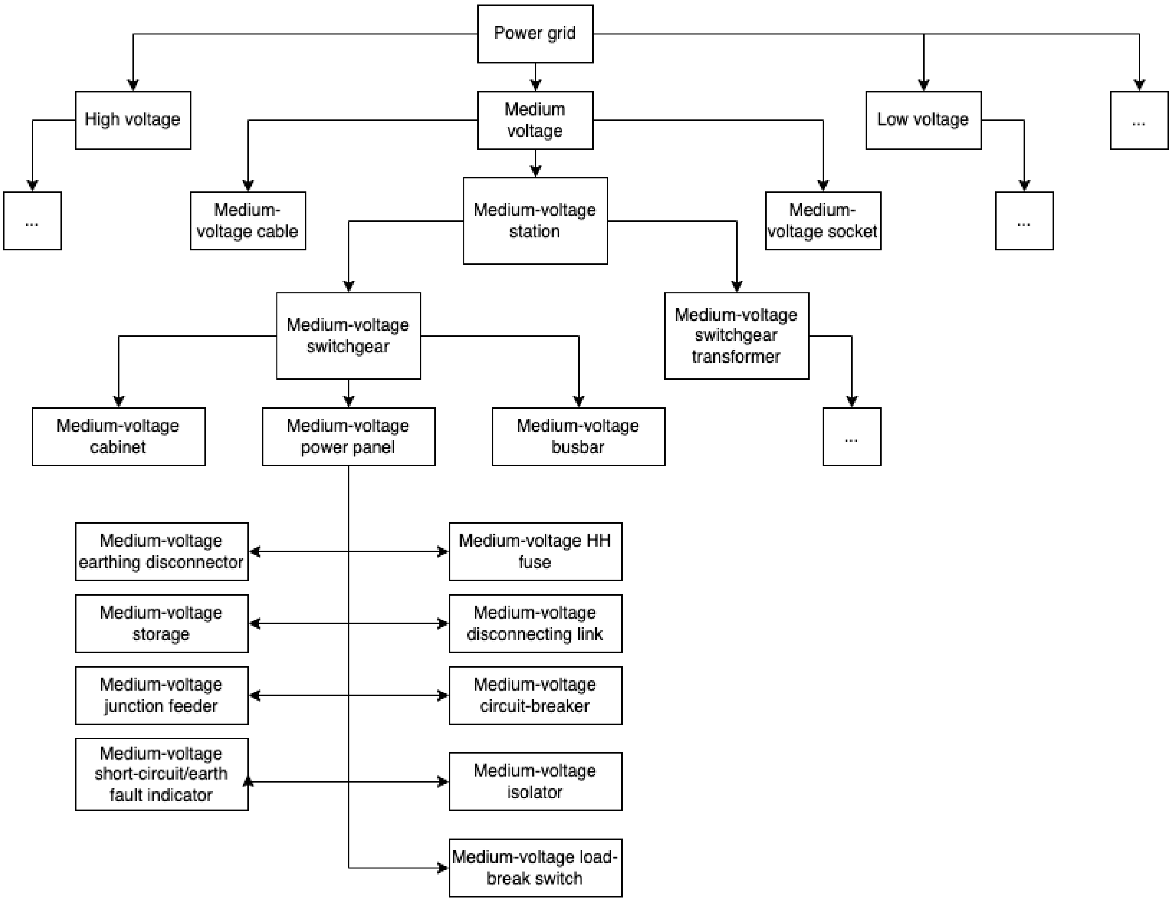

Figure 4. Anonymized Class Modell highlighting the Power Grid

Furthermore, this data also aids in gathering tacit employee knowledge as they utilize these lists daily. The employee knowledge was gathered in a workshop. We formed five groups, with one researcher and two employees each. The knowledge was gathered in three different ways: top-down and bottom-up in the hierarchy of assets, or via a process-driven approach. These approaches were combined with storytelling and after-action reviews to increase the depth of the externalized tacit knowledge. Table 1 shows the experience of the employees and their groups. By combining and comparing these knowledge artifacts within the groups and across the groups, we were able to validate the knowledge gathered in our workshops. Finally, these approaches resulted in 169 knowledge artifacts across 23 different related assets within the knowledge collection from the employee knowledge-gathering activity. Since this iteration only served as an initial demonstration of the method, further iterations across other high-level power grid components are yet to be made. The established knowledge could, therefore, be structured and used as the base for the knowledge collection to provide tacit knowledge for an LLM model as part of an assistance system.

Group	1		2		3		4		5	
Employee	1	2	3	4	5	6	7	8	9	10
Experience (in years)	8	9	10	4	5	8	14	8	4	4

Table 1. Employee Experience and workshop groups during the knowledge-gathering activity

Figure 5. Exemplary Buffer Analysis of Environmental Influences on the MV Station ST7060XX

Based on this instance model, we were also able to gather the areas in which the relevant power grid stations were positioned and follow the activities associated with geodata. In an exemplary iteration, the Geographic Information System (GIS) ArcGIS was used to locate a power station and map nearby pollutants and environmental influences for this station. Figure 5 shows different influences on the MV station, like a forest and a rapeseed field. Forest wildlife and increased humidity, are within the area and could cause disruptions to the power stations. Further, pollen and dust from the rapeseed field might be carried into the facility by the wind. From these areas, approximations for the spread of these environmental influences could be made. The areas are thus expanded by zones that could potentially be reached by these influences. The data is combined and tied to the areas as part of the managing location data within the GIS. As part of the structuring phase the data is then clustered as visualized in Table 2.

Area	**Coordinates**	**Surrounding**	**Biosphere**	**Weather**
Rapeseed field	X,71695N, X,64240E; …	Pollutants		
Forest	X,71740N, X,63711E; …		Forest wildlife	Increased humidity

Table 2. Exemplary Selection of the Geodata Collection.

Table 2 and Table 3 thus show an exemplary part of the knowledge collection based on the asset and geodata. The tables present the first iteration of the knowledge collection with results from the demonstration and include the gathered knowledge artefacts. Both collections can be combined by gathering coordinates from the assets and looking for the area in the geodata collection. For example, ‘MV Station ST7060XX’ is located at coordinates within the Polygon A area. Within this area, mostly rainy weather and increased humidity occur; therefore, a different degradation of the assets within the station when compared to other stations should be expected. Furthermore, it is known that the station includes multiple switchgears. During the maintenance process, when the employee is preparing the vehicle, it should be made sure that enough cleaning products are packed highlighting explizit knowledge gathered from source systems.

Asset	**Process**	**Activity**	**Knowledge Artefact**
MV Station ST7060XX	Maintenance	Prepare maintenance	X,717436N, X,636726E
MV Station ST7060XX	Maintenance	Prepare vehicle	Large station with many switchgears, might require more cleaning products

Table 3. Exemplary Section of the First Knowledge Collection Iteration.

4.3 Evaluation

The presented IRIS method offers the ability to combine systems data and employee knowledge to structure knowledge. From the demonstration of the first iteration, the main adjustment to the method regards the retrieval and investigation phase. During the workshop, top-down, bottom-up, and process-driven approaches were tested. In the workshop, the participants naturally gravitated toward the process-driven approach—aligning with their intuitive communication style—which clearly demonstrated that describing specific cases is the most effective for externalizing employee knowledge. Therefore, going forward, the recommendation for the knowledge-gathering stages is to emphasize a process-driven approach in the retrieval and investigation phase. Another change made was during the structuring of knowledge for the knowledge collection. For the assistance system to differentiate between class and instant specific knowledge, an indicator had to be made to the knowledge collection table. Table 4 now highlights the final knowledge collection. Knowledge can apply to all classes, like the knowledge artifact “schedule activity *switch call* one day ahead at the latest”, thus externalizing tacit knowledge. Instance specific knowledge, like the previously introduced additional need for cleaning products, is now firstly categorized as the class ‘MV Station’ and then appended to the instance ‘MV Station ST7060XX’. This should allow the assistance system to distinguish between instance-specific knowledge and general class knowledge. Additionally, since the instances are derived from the technical names, it can easily be mapped to the data in the ERP system, creating a combined knowledge base.

Class	**Instance**	**Process**	**Activity**	**Knowledge Artefact**
MV Station	-	Maintenance	Switch call	Schedule one day ahead at the latest
MV Station	MV Station ST7060XX	Maintenance	Prepare vehicle	Large station with many switchgears, might require more cleaning products

Table 4. Example of the Final Knowledge Collection

Due to this combination, from a practical standpoint duplicate knowledge does not matter. The GenAI behind the assistance system should be able to cite multiple sources for the knowledge presented. Therefore, clashes between instance and class knowledge should not be an issue during the application of the method and later on in the usage of the assistance system.

In general, the presented method for gathering knowledge in maintenance processes worked well. Practitioners highlighted how quickly and precise knowledge artefacts can be gathered “by just talking about them” as per the training supervisor. While this demonstration mainly focused on the activities within the method and its outcomes, we aim to evaluate the method with the complete assistance system at a later point. To do so, we plan to evaluate the usage of the assistance system with the results from the knowledge collection and without them—thus, solely focusing on systems data and already externalized knowledge. Only at that point will it be possible to fully capture the improvements made to the expert assistance system by utilizing tacit knowledge in addition to the system's data and already externalized knowledge.

Summarizing our findings, we introduce the IRIS method for systematically collecting and structuring domain-specific knowledge to support the development of assistance systems in maintenance. Through a DSR study, we integrated explicit and tacit knowledge, particularly from employees, as well as location data to establish a comprehensive knowledge base for GenAI-powered assistance systems. We demonstrated our method with a distribution grid operator, successfully capturing 169 knowledge artifacts in a single workshop, demonstrating and evaluating the method.

5 Discussion

Gathering knowledge, especially in its tacit form, is generally perceived as difficult. The latest changes in GenAI significantly changed the way this can be done. We agree with Santoro et al. (2018) that such disruptive technologies impact KM approaches. Their description of the technological progress among economic players highlights the need for a multi-disciplined approach to the topic of KM. Therefore, utilizing existing research from the domains of business process management and data science, especially in a GenAI setting, is useful to consider. Especially as the need for context and contextual characteristics in business process management should be considered (vom Brocke et al., 2016). Therefore, creating suitable methods that not only take the processes into consideration but also the people fulfilling them. Establishing a combination of explicit and tacit knowledge must be done (Santoro et al., 2018). Furthermore, selecting a process-driven approach in combination with a matching knowledge-gathering approach (Seghroucheni et al., 2023) seems sensible. Even if this leads to a switch from an asset-driven outline to a process-driven approach in the investigation phase.

Even though GIS improve decision-making and enable innovative applications especially in smart grid and IoT scenarios (Ashkezari et al., 2018; Cao & Wachowicz, 2019; zur Heiden et al., 2022), they are still an underrepresented topic in IS research (Priefer, 2023). Integrating GIS into maintenance processes is particularly vital for distribution grid providers, as it allows for a location-based approach to managing asset conditions and predicting maintenance needs. Domain-specific data is essential for effective decision-making in specialized areas and geospatial data should be incorporated into the knowledge management process (zur Heiden et al., 2022). By linking transactional and geospatial data, GIS enable a more dynamic interaction between data collection and knowledge generation, facilitating a comprehensive view of maintenance needs tied to geographic variables. This dual approach not only supports the tailored gathering of geospatial data but also offers a transferable framework that other sectors could adapt. By replacing GIS elements with relatable systems, industries outside the energy sector could similarly enhance their processes. For instance, the impact of IoT integration on innovation projects (Santoro et al., 2018). Our paper thus contributes to the literature by demonstrating how GIS-enhanced methodologies can improve maintenance efficiency and broaden the scope of data-driven knowledge management in IS research.

Given this wide range of data basis, we expand existing knowledge gathering methods in multiple dimensions. While previous research utilized maintenance manuals and combined them with named entities recognition to establish a knowledge graph (Z. Liu & Lu, 2024), IRIS takes on these concepts and further expands them by adding tacit knowledge. While LLMs excel at processing unstructured text, structured data offers advantages in domain-specific applications. In maintenance contexts, where AI-driven recommendations need to be precise and context-aware, structuring ensures that models can reliably associate knowledge with specific assets, processes, and environmental factors. Although systems for building maintenance based on a knowledge base are not new, there is a lack of professional knowledge on a larger scale, e.g., for various construction operations (Motawa & Almarshad, 2013). The integration of domain-specific systems, like the implementation of geodata in the IRIS method, addresses this gap. While our method uses system data to create hierarchical instances and class models like in previous research that establishes conceptual graphs to model past experiences and extract rules from business logic (Potes Ruiz et al., 2014), IRIS places less emphasis on defining rules. This is due to the fact that business rules should be represented within the business processes (Dumas et al., 2018a). Additionally, in a maintenance-specific context, manuals and checklists exist to specify boundaries within the activities.

Thus, our presented method combines the research topics of assistance systems and KM, focusing on domain-specific data and process data. As our method is the first to include tacit knowledge and its collection as an input for a GenAI system, researchers and practitioners can, therefore, modify the method to suit the specific data sources they encounter. Within the maintenance domain, researchers can use the method to establish a knowledge base for their assistance systems and add specific modifications for contexts not relevant to maintenance tasks in power grids. Outside of the maintenance domain, researchers can substitute the GIS-specific part of our method with systems relevant to their domain—abstracting our method to a context beyond maintenance. Practitioners can use the method to collect data and knowledge from their systems and employees, i.e., tacit knowledge and systems data. Building on this overview, they can then train their GenAI to build a system that is prone to less hallucinations and outputs results that match the peculiarities of the context much better.

6 Conclusion

We utilized DSR to design a method for the gathering of knowledge across systems, processes and employees that can be used to train GenAI applications. The method has been demonstrated at a regional distribution grid provider. By utilizing the underlying ERP systems, it was possible to quickly gather enough data for instance and class models that almost match reality. Over the course of three workshops for gathering process insights and one workshop with ten experts to gather employee knowledge, it was possible to gather and structure 169 tacit knowledge artifacts. These artifacts are further expanded by a growing collection of location data and source systems information like maintenance manuals and checklists.

While these results are impressive, they do come with some limitations. The demonstration was conducted at one medium-sized distribution grid provider. While it can be assumed that the method works well in similar settings, larger or smaller distribution grid providers with a different IT infrastructure might achieve different results. While it has been confirmed that the gathering of knowledge works with the method, developing the expert assistance system and thus achieving a complete evaluation is still pending, highlighting a major limitation. Future research should also validate whether the IRIS method is efficient and effective in further contexts, maybe even extending to non-maintenance domains.

Nevertheless, the presented method allows for a multi-layered approach to gathering and structuring knowledge in maintenance processes. Our next research step is the development of an assistance system for maintenance processes at distribution grid providers, i.e., use our method as an input device for building a GenAI assistance system in maintenance. To do so, GenAI must be able to incorporate the data from the knowledge collection and dynamically add or archive outdated knowledge. Given the current status, it is expected that the created knowledge base will serve as an initial spark to start such a system. Advanced GenAI techniques like human-in-the-loop systems can feed from these aspects but might require additional alignment and feedback systems to function properly. Future research should also target the application of our method in neighboring and different domains to identify peculiarities of maintenance and other contexts. Other avenues for future research are the application of our method in different domains and comparing trained GenAI application to verify the usefulness of IRIS.

7 Acknowlegements:

This article was written as part of the research and development project “AProSys - AI-supported assistance and forecasting systems for sustainable use in intelligent distribution grid technology”. The project is funded by the German Federal Ministry for Economic Affairs and Climate Protection (BMWK) under funding code 03EI60090E.

References

Ashkezari, A. D., Hosseinzadeh, N., Chebli, A., & Albadi, M. (2018). Development of an enterprise Geographic Information System (GIS) integrated with smart grid. Sustainable Energy, Grids and Networks, 14, 25–34. https://doi.org/10.1016/j.segan.2018.02.001

Azevedo, N., Aquino, G., Nascimento, L., Camelo, L., Figueira, T., Oliveira, J., Figueiredo, I., Printes, A., Torné, I., & Figueiredo, C. (2023). A Novel Methodology for Developing Troubleshooting Chatbots Applied to ATM Technical Maintenance Support. Applied Sciences, 13(11), 6777. https://doi.org/10.3390/app13116777

Bannat, A. (2014). Ein Assistenzsystem zur digitalen Werker-Unterstützung in der industriellen Produktion [An assistance system for digital worker support in industrial production]. Technische Universität München. https://mediatum.ub.tum.de/1206607

BMWK. (2024, November 12). Das deutsche Stromnetz. BMWI. https://www.bmwk.de/Redaktion/DE/Infografiken/Energie/abbildung-das-deutsche-stromnetz.html

Cao, H., & Wachowicz, M. (2019). The design of an IoT-GIS platform for performing automated analytical tasks. Computers, Environment and Urban Systems, 74, 23–40. https://doi.org/10.1016/j.compenvurbsys.2018.11.004

Czako, V. (2020). Employment in the energy sector: Status report 2020. JRC science for policy report. Publications Office of the European Union. https://doi.org/10.2760/95180

Di Ciccio, C., Marrella, A., & Russo, A. (2015). Knowledge-Intensive Processes: Characteristics, Requirements and Analysis of Contemporary Approaches. Journal on Data Semantics, 4(1), 29–57. https://doi.org/10.1007/s13740-014-0038-4

Dueker, K. J., & Kjerne, D. (1989). Multipurpose cadastre: Terms and definitions. American Society for Photogrammetry and Remote Sensing: American Congress on Surveying and Mapping.

Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2018). Introduction to Business Process Management. In M. Dumas, M. La Rosa, J. Mendling, & H. A. Reijers (Eds.), Fundamentals of business process management (Second edition, pp. 1–33). Springer. https://doi.org/10.1007/978-3-662-56509-4_1

Ehlers, M., & Schiewe, J. (2012). Geoinformatik [Geoinformatics]. Raumforschung Und Raumordnung | Spatial Research and Planning, 70(6), 557–559. https://doi.org/10.1007/s13147-012-0183-x

Elsahlamy, E., Eshra, A., Eshra, N., & El-Fishawy, N. (2021). Empowering GIS with Big Data: A review of recent advances. In Iceem2021: 2nd IEEE International Conference on Electronic Engineering, Menoufia University, Egypt : Faculty of Electronic Engineering, Menouf, Egypt, 3-4 July 2021 (pp. 1–7). IEEE. https://doi.org/10.1109/ICEEM52022.2021.9480634

Erden, Z., Krogh, G. von, & Nonaka, I. (2008). The quality of group tacit knowledge. The Journal of Strategic Information Systems, 17(1), 4–18. https://doi.org/10.1016/j.jsis.2008.02.002

Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), Official Journal of the European Union 1 (2007).

Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. Business & Information Systems Engineering, 66(1), 111–126. https://doi.org/10.1007/s12599-023-00834-7

Fill, H.-G. (2024). Spatial Conceptual Modeling: Anchoring Knowledge in the Real World. Advance online publication. https://doi.org/10.48550/arXiv.2407.17259

Fill, H.-G., Härer, F., Vasic, I., Borcard, D., Reitemeyer, B., Muff, F., Curty, S., & Bühlmann, M. (2024). CMAG: A Framework for Conceptual Model Augmented Generative Artificial Intelligence. ER2024: Companion Proceedings of the 43rd International Conference on Conceptual Modeling: ER Forum, Special. https://ceur-ws.org/vol-3849/forum5.pdf

García-Fernández, M. (2015). How to measure knowledge management: Dimensions and model. VINE, 45(1), 107–125. https://doi.org/10.1108/VINE-10-2013-0063

Girard, J [John], & Girard, J [JoAnn] (2015). Defining knowledge management: Toward an applied compendium. Online Journal of Applied Knowledge Management, 3(1).

Gouvêa, M. T. A., Santoro, F. M., Cappelli, C., & Pimentel, M. (2016). Knowledge Management in Distance Education. In ACM Digital Library, Proceedings of the The 11th International Knowledge Management in Organizations Conference on The changing face of Knowledge Management Impacting Society (pp. 1–8). ACM. https://doi.org/10.1145/2925995.2926000

Hevner, March, Park, & Ram (2004). Design Science in Information Systems Research. MIS Quarterly, 28(1), 75. https://doi.org/10.2307/25148625

Keenan, P. (1997). Geographic Information SystemsTheir contribution to the IS mainstream. AMCIS 1997 Proceedings, Article 127.

Link, M., Dukino, C., Ganz, W., Hamann, K., & Schnalzer, K. (2020). The Use of AI-Based Assistance Systems in the Service Sector: Opportunities, Challenges and Applications. In I. L. Nunes (Ed.), Advances in Intelligent Systems and Computing. Advances in Human Factors and Systems Interaction (Vol. 1207, pp. 10–16). Springer International Publishing. https://doi.org/10.1007/978-3-030-51369-6_2

Liu, X., Vatn, J., Yin, S., & Maithani, V. (2023). Performance of ChatGPT on CMRP: Potential for Assisting Maintenance and Reliability Professionals Using Large Language Models. In Iecon 2023 - 49th Annual Conference of the IEEE Industrial Electronics Society: 16-19 Oct. 2023 : Conference location: Singapore, Singapore (pp. 1–7). IEEE. https://doi.org/10.1109/IECON51785.2023.10311736

Liu, Z., & Lu, Y. (2024). A task-centric knowledge graph construction method based on multi-modal representation learning for industrial maintenance automation. Engineering Reports, Article e12952. Advance online publication. https://doi.org/10.1002/eng2.12952

Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S., & Söllner, M. (2019). AI-Based Digital Assistants. Business & Information Systems Engineering, 61(4), 535–544. https://doi.org/10.1007/s12599-019-00600-8

Maedche, A., Morana, S., Schacht, S., Werth, D., & Krumeich, J. (2016). Advanced User Assistance Systems. Business & Information Systems Engineering, 58(5), 367–370. https://doi.org/10.1007/s12599-016-0444-2

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. Decision Support Systems, 15(4), 251–266. https://doi.org/10.1016/0167-9236(94)00041-2

McAdam, R., Mason, B., & McCrory, J. (2007). Exploring the dichotomies within the tacit knowledge literature: towards a process of tacit knowing in organizations. Journal of Knowledge Management, 11(2), 43–59. https://doi.org/10.1108/13673270710738906

Morana, S., Gerards, T., & Mädche, A. (2015). ITSM ProcessGuide – A Process Guidance System for IT Service Management. In B. Donnellan, M. Helfert, J. Kenneally, D. VanderMeer, M. Rothenberger, & R. Winter (Eds.), Lecture Notes in Computer Science. New Horizons in Design Science: Broadening the Research Agenda (Vol. 9073, pp. 406–410). Springer International Publishing. https://doi.org/10.1007/978-3-319-18714-3_32

Motawa, I., & Almarshad, A. (2013). A knowledge-based BIM system for building maintenance. Automation in Construction, 29, 173–182. https://doi.org/10.1016/j.autcon.2012.09.008

Murphy, L. D. (1996). Competing in space: The strategic roles of geographic information systems. AMCIS 1996 Proceedings, Article 96.

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. Organization Science, 5(1), 14–37. https://doi.org/10.1287/orsc.5.1.14

O'Dell, C., O'Dell, C. S., & Grayson, C. J. (1998). If only we knew what we know: The transfer of internal knowledge and best practice. Free Press. https://doi.org/10.2307/41165948

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. Journal of Management Information Systems, 24(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Peissner, M., Kötter, F., & Zaiser, H. (2019). Künstliche Intelligenz – Anwendungsperspektiven für Arbeit und Qualifizierung [Artificial intelligence - application perspectives for work and qualification]. Berufsbildung Wissenschaft Prax, 48(3), 9–13.

Perri, L. (2024). 3 Bold and Actionable Predictions for the Future of GenAI. Gartner. https://www.gartner.com/en/articles/3-bold-and-actionable-predictions-for-the-future-of-genai

Philipson, S., & Kjellström, E. (2020). When objects are talking: How tacit knowing becomes explicit knowledge. Journal of Small Business Strategy (Archive Only), 30(1), 68–82. http://libjournals.mtsu.edu/index.php/jsbs/article/view/1666

Piñeiro-Martín, A., García-Mateo, C., Docío-Fernández, L., & Del López-Pérez, M. C. (2023). Ethical Challenges in the Development of Virtual Assistants Powered by Large Language Models. Electronics, 12(14), 3170. https://doi.org/10.3390/electronics12143170

Polanyi, M. (2009). The tacit dimension. Reproduction en fac-similé. University of Chicago Press.

Potes Ruiz, P., Kamsu Foguem, B., & Grabot, B. (2014). Generating knowledge in maintenance from Experience Feedback. Knowledge-Based Systems, 68, 4–20. https://doi.org/10.1016/j.knosys.2014.02.002

Priefer, J. (2023). Geographic Information Systems in Information Systems Research-Review and Research Prospects. Proceedings of the 56th Hawaii International Conference on System Sciences, 5048–5058.

Refaiy, M., & Labib, A. (2009). The effect of applying tacit knowledge on maintenance performance: an empirical study of the energy sector in the UK and Arab countries. Knowledge Management Research & Practice, 7(3), 277–288. https://doi.org/10.1057/kmrp.2009.11

Sabri, M. O. (2023). Mapping Knowledge-Life Cycle to Process Architecture. European Conference on Knowledge Management, 24(2), 1173–1177. https://doi.org/10.34190/eckm.24.2.1249

Santoro, G., Vrontis, D., Thrassou, A., & Dezi, L. (2018). The Internet of Things: Building a knowledge management system for open innovation and knowledge management capacity. Technological Forecasting and Social Change, 136, 347–354. https://doi.org/10.1016/j.techfore.2017.02.034

Schulz, M., Neuhaus, U., Kaufmann, J., Badura, D., Kerzel, U., Welter, F., Prothmann, M., Kühnel, S., Passlick, J., Rissler, R., Badewitz, W., Dann, D., Gröschel, A., Kloker, S., Alekozai, E. M., Felderer, M., Lanquillon, C., Brauner, D., Gölzer, P., . . . Gehrke, N. (2021). DASC-PM v1.0 - Ein Vorgehensmodell für Data-Science-Projekte [A process model for data science projects]. https://doi.org/10.25673/32872.2

Seghroucheni, O. Z., Achhab, M. A., & Lazaar, M. (2023). Systematic Review on the Conversion of Tacit Knowledge. In 2023 7th IEEE Congress on Information Science and Technology (CiSt) (pp. 123–128). IEEE. https://doi.org/10.1109/cist56084.2023.10409884

Seufert, S., & Meier, C. (2023). Hybride Intelligenz: Zusammenarbeit mit KI-Assistenzsystemen in wissensintensiven Bereichen [Hybrid intelligence: collaboration with AI assistance systems in knowledge-intensive areas]. HMD Praxis Der Wirtschaftsinformatik, 60(6), 1194–1209. https://doi.org/10.1365/s40702-023-01012-9

Shin, W., Han, J., & Rhee, W. (2021). AI-assistance for predictive maintenance of renewable energy systems. Energy, 221, 119775. https://doi.org/10.1016/j.energy.2021.119775

Smith, O., Cattell, O., Farcot, E., O'dea, R. D., & Hopcraft, K. I. (2022, March 2). The effect of renewable energy incorporation on power grid stability and resilience. American Association for the Advancement of Science. https://www.science.org/doi/10.1126/sciadv.abj6734

Stodden, V. (2020). The data science life cycle: a disciplined approach to advancing data science as a science. Communications of the ACM, 63(7), 58–66. https://doi.org/10.1145/3360646

Stokes, M. J. (2001). Managing Engineering Knowledge: MOKA-Methodology for Knowledge Based Engineering Applications.

Szelągowski, M. (2021). The knowledge and process dimensions. VINE Journal of Information and Knowledge Management Systems, 51(2), 271–287. https://doi.org/10.1108/VJIKMS-09-2019-0150

vom Brocke, J., Zelt, S., & Schmiedel, T. (2016). On the role of context in business process management. International Journal of Information Management, 36(3), 486–495. https://doi.org/10.1016/j.ijinfomgt.2015.10.002

Werrmann, J., & Krämer, B. J. (2018). Knowledge Management through Ontology-Driven Integration of Disparate Knowledge Sources. Journal of Integrated Design and Process Science, 21(4), 25–39. https://doi.org/10.3233/jid-2017-0017

Wiig, K. M. (1998). Knowledge management foundations: Thinking about thinking; how people and organizations create, represent, and use knowledge (5th ed.). Knowledge management / Karl M. Wiig: Vol. 1. Schema Press.

Zhang, Z., & Yu, L. (2014). Acedmic Hot-Spot Analysis on Information System Based on the Co-Term Network. PACIS 2014 Proceedings, 2014, Article 109.

zur Heiden, P., Priefer, J., & Beverungen, D. (2022). Utilizing Geographic Information Systems for Condition-Based Maintenance on the Energy Distribution Grid. Proceedings of the 55th Hawaii International Conference on System Sciences (HICSS), 5672–5681.

zur Heiden, P., Priefer, J., & Beverungen, D. (2023). Location-Based Service and Location-Contextualizing Service: Conceptualizing the Co-creation of Value with Location Information. *Hawaii International Conference on System Sciences 2023 (HICSS-56)*. https://aisel.aisnet.org/hicss-56/li/research/2