Tacit Knowledge Elicitation Process for Industry 4.0

Enzo Fenoglio (✉ e.fenoglio@ucl.ac.uk)

University College London

Emre Kazim

University College London

Hugo Latapie

Emerging Technologies & Incubation, Cisco Systems Inc

Adriano Koshiyama

University College London

Research Article

Keywords: concept maps, knowledge graph, ontology, tacit knowledge, knowledge management, AI ethics

Posted Date: December 8th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1146087/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Discover Artificial Intelligence on March 10th, 2022. See the published version at <https://doi.org/10.1007/s44163-022-00020-w>.

Tacit Knowledge Elicitation Process for Industry 4.0

Enzo Fenoglio^{1,3*}, Emre Kazim^{1,2}, Hugo Latapie³
and Adriano Koshiyama^{1,2}

^{1*}Department of Computer Science, University College London,
Glower St, London, WC1E 6BT, UK.

²Holistic AI, 18 Soho Square, London, W1D 3QL, UK.

³Emerging Technologies & Incubation, Cisco Systems Inc, 170
West Tasman Dr., San Jose, 95134, CA, USA.

*Corresponding author(s). E-mail(s): e.fenoglio@ucl.ac.uk;

Contributing authors: emre.kazim@holisticai.com;

hlatapie@cisco.com; adriano.koshiyama@holisticai.com;

Abstract

Manufacturers migrate their processes to Industry 4.0 that includes new technologies for improving productivity and efficiency of operations. One of the issues is capturing, recreating, and documenting the tacit knowledge of the aging workers. However, there are no systematic procedures to incorporate this knowledge into Enterprise Resource Planning systems for maintaining a competitive advantage. This paper describes a solution proposal for a tacit knowledge elicitation process for capturing operational best practices of experienced workers in industrial domains for on-site training. We use Concept Maps mining and domain ontologies to discover and integrate information into a Knowledge Graph in a role game played by human and virtual agents. Ethical and societal concerns are discussed as well.

Keywords: concept maps, knowledge graph, ontology, tacit knowledge, knowledge management, AI ethics.

1 Introduction

Manufacturers migrate their processes to Industry 4.0 innovative practices, which include the adoption of recent technologies for improving productivity and efficiency of operations through visibility and analytics [1]. One of the major critical issues manufacturers face is capturing, recreating, and documenting the experience of the aging workforce, the so-called *tribal knowledge* before they change role or leave the company — This paper uses indifferently the terms of *tribal knowledge*, *tacit knowledge*, and *implicit knowledge*, although a subtle difference between *tacit* and *implicit* knowledge and the term *tribal knowledge* is jargon.

Tribal knowledge is a term widely employed in the industry to denote critical knowledge obtained by some senior staff subject matter experts (SME) who have gained deep expertise on types of equipment, a device, or a method. It is associated with action since it reflects *understanding how* more than *knowing what* [2]. The **Six Sigma Business Dictionary** describes tribal knowledge as “*any unwritten information that is not commonly known by others within a company. This term is used most when referencing information that may need to be known by others to produce a quality product or service*”. The tribal knowledge represents how people act unconsciously and intuitively. It is always tacit and never expressed [3] or not easily expressible since *we can know more than we can tell* [4]. Tribal knowledge is a subset of the institutional knowledge, which comprises all documented and undocumented knowledge in an organization [5]. It brings decades of hands-on experience acquired without direct instruction or self-study or help from others. In this sense, it belongs to the company, but it is *stored within the heads* of the experienced workforce never transformed into the company knowledge base, and quantifiable only indirectly as a loss when senior workers leave and cannot get replaced by apprentices with comparable performance skills. The process of extracting information *out of the head* of an expert is not new. It was part of the development process for expert systems. The method consisted of two phases: (i) knowledge elicitation, where the knowledge was extracted by the expert, and (ii) knowledge representation, where the knowledge was stored in a database. The techniques utilized in the past were somehow inefficient and based on direct interviews to develop rule-based systems without the support of contemporary cognitive psychology [6] or knowledge management principles [5]. This paper follows the same bi-partition, but we describe a knowledge elicitation process conducted by human and virtual agents in a role game where the facts discovered will be represented in a Knowledge Graph (KG). The synergy between SMEs and knowledge engineers from one side and the cognitive system from the other will lead to knowledge creation through constructive learning, reflective ability, active interaction, and collaboration of human and virtual participants. The conversion of tacit knowledge into organizational knowledge will be promoted by (i) the application of Concept Maps (CMs) mining for concepts visualization [7], (ii) the application of domain ontology on a KG for knowledge

representation and automatic knowledge generation [8], and (iii) the application of logical and semantic reasoning to infer new knowledge in a continuous learning process for knowledge retrieval [9].

The rest of the paper is organized as follows: In Section 2, we present some of the methods proposed for capturing tacit knowledge. In Section 3, we provide a functional description of our cognitive framework for capturing tacit knowledge. In Section 4, we describe in detail the knowledge elicitation process using a game role approach. In Section 5, we discuss the societal and ethical concerns of the system proposed. In Section 6, we draw some conclusions. Future works are discussed in Section 7.

2 Related Work

Tacit knowledge is the work-related practical knowledge learned informally on the job by workers. However, the importance of tacit knowledge is not systematically recognized by most companies, probably because the workforce age segmentation has never impacted companies as faced nowadays. According to the **US Bureau of Labor Statistics**¹, by 2029, around 25% of the workforce in the industry will be aged above 55 years and retire in the following years. As many qualified people retire, a wealth of information related to *best practices* and *efficient operations* is being inevitably lost. It will become increasingly difficult to find experts with 25 or more years in the workforce who know how to fix a critical problem — assuming that it is non-trivial to transfer experience across jobs. Eventually, the better an organization can elicit tacit knowledge from its employees and share it across the organization, the more innovative it can be [10].

Many companies do not have any systematic approach to collect and incorporate the field experience of experienced workers and assess the role of tacit knowledge management on the success of Enterprise Resource Planning (ERP) systems implementation [11]. At best, the loss of tacit knowledge is mitigated by informal procedures [6]. The main approaches can be listed as follows:

1. **Internal reports:** For most companies, the standard best practice for recording tacit knowledge is via internal reports. However, this procedure, if not enforced, is ineffective and difficult to replicate. A critical aspect of this approach is related to workers' communication skills. SMEs who have previous experience expressing their expertise are far more able to share their knowledge than others. On the contrary, workers who lack minimal communication skills with similar domain expertise cannot externalize their knowledge correctly to others. A better approach is described in [12] but without considering tacit knowledge.
2. **One-on-one mentoring:** Most companies prefer to organize a worker development program in the form of one-on-one mentoring [13], where a new apprentice shadows some veteran engineer, technician, or senior worker in the first 6 or 12 months on the job. The tacit knowledge is usually passed

¹www.bls.gov

on during on-site training. However, due to the urgency of the assigned tasks, senior engineers typically do not have the time, energy, and motivation to carry out maintenance inspection and fault diagnosis and thoroughly explain *how they function while working* [6] to supervisors.

3. **On-site training:** Some companies prefer to provide on-site training to young engineers in the belief that they may better absorb and grasp domain knowledge from veteran engineers after they have a mental picture of the overall system or problem they are working to. However, most of the time, mentoring focuses on resolving an individual issue but fails to help trace the root cause of the problem in similar systems. It may often cause the inexperienced workers to jump directly to problem-solving using inefficient (sometimes dangerous) shortcuts for trial and error or opting for unwarranted large-scale parts or components replacement when charged with similar jobs.
4. **Rules collector system:** This method enables a process to capture an individual's expertise in a formalized manner updating hand-crafted rules and knowledge databases with information retrieved from the interaction with workers. It is the old approach from the time of expert systems. Unfortunately, rule-based systems are brittle and difficult to maintain. The rules captured can be contradictory and non-sufficient to ensure the consistency of the knowledge base. In our approach, we avoid the limitations of the rules collector system by reasoning in domain ontology to discover new knowledge.
5. **Education Domain:** The tacit knowledge elicitation process that we describe for constructing the *operational knowledge graph* in industrial domains (Section 4) is similar to the automatic construction of KGs in education domains. For example, in [14], the authors introduce the *KnowEdu* system that supports personalized teaching services and adaptive learning solutions using pedagogical data in textual format. However, in their approach, the structure of the KG is induced from the data, while our approach is modeled from an ontology.

3 The Cognitive Framework Functional Architecture

In this section, we describe a cognitive pipeline for the conversion of tacit knowledge into explicit knowledge. The mix of explicit and tacit knowledge allows organizations to make sense of their environment [15]. The continuous social interaction of tacit and explicit knowledge through dialogue and debate creates the institutional knowledge maintained on the internal technical documentation of the company. The implicit knowledge represents the industrial culture, the occupational traditions, and the cultural values of the workforce that uses, develops, administers, and operates the technology in the working environment [16]. In this sense, it is very reductive to think that the overall knowledge in a company belongs exclusively to one particular group of experts

or a single individual as competence, skills, or know-how. It is far better to consider this knowledge as the result of social accomplishments of constructing and reconstructing new knowledge as the ongoing product of practices that engage employees, promote collaboration, and expand the knowledge transformation towards a stronger company culture where background knowledge and implicit cognitive rules have a fundamental relevance. Tacit knowledge is a social and not an individual attribute [17], and the different forms of tacit knowledge [18] go beyond the reductive form of tacit knowledge as know-how. If we want to efficiently translate tacit knowledge into explicit knowledge, we shall also emulate the social working environment in which the workforce is acting. Insofar we assume the applicability of the *Nonaka-Takeuchi model* [19] that postulates how knowledge in an organization is created through continuous social interaction of tacit and explicit knowledge, coupled with the four sequential modes of knowledge conversion of the SECI model: Socialization (from tacit to tacit), Externalisation (from tacit to explicit), Combination (from explicit to explicit), and Internalization (from explicit to tacit). It is outside the scope of this paper to discuss the limitations and the criticism raised by the SECI knowledge creation process. More detailed information on these aspects can be found in [20–23].

For the objectives of this paper, the SECI spiral model of knowledge creation perfectly justifies the role game we propose to convert tacit knowledge into explicit knowledge with human and virtual agents interacting in the process of knowledge transformation. It is worthwhile noting how the inability to formalize directly tacit knowledge does not exclude the possibility that a virtual agent (a computer system) might perform the same tasks using alternative representations or that tacit knowledge cannot be transferred to a machine [24, 25]. Nevertheless, no automatic cognitive system can capture the tacit knowledge or any other genuine human activity with a purely algorithmic approach —from workers’ heads to corporate databases— without considering the social, cultural, legal, sociological contexts of the data collection and representation. The supervision of a human agent in the cooperative game described in Section 4 will ensure that non-algorithmic qualitative factors that require human appreciation may be considered during the knowledge transformation process, complementing the neuro-symbolic pipeline used for capturing the algorithmic quantitative factors.

3.1 The Neurosymbolic Architecture

The architecture of the cognitive framework for tacit knowledge elicitation can be described as a hybrid neurosymbolic system where a neural network focused on sub-symbolic tasks interacts with a symbolic system — it can be classified as **Neuro;Symbolic** (Type 3) [26].

The functional architecture of the cognitive framework is presented in Fig. 1. It includes three main layers: (i) sub-symbolic layer, (ii) conceptual layer, (iii) symbolic layer. The main difference with other neurosymbolic architectures is

Fig. 1 Cognitive framework Functional Architecture for tacit knowledge elicitation.

that the symbolic and the sub-symbolic layers are not coupled directly but interact through an intermediate layer (the conceptual layer).

3.1.1 Sub-symbolic Layer

The sub-symbolic layer processes heterogeneous input data collected from the internal technical documentation, operation and maintenance (O&M) manuals, troubleshooting manuals, technical documentation and reports, from different structured or unstructured sources. Other input data types, such as video, and audio have not been scoped in this study but can be recorded for archival purposes and documentation.

Raw Data Processing

One of the advantages of using input in textual format is the great availability of mature Natural Language Processing (NLP) methods. Actually, the raw input textual data is stored in files and processed using standard data mining tools for automatic knowledge extraction from text, such as **Stanford NLP software**² and **Apache OpenNLP**³. The objective is the extraction of textual information from the documents for syntactic analysis, Named Entity Recognition (NER), Relation Extraction (RE), and Query Generation (QG).

3.1.2 Conceptual Layer

The conceptual layer bridges the gap between symbolic and sub-symbolic representation with an intermediate layer that shares knowledge structures using a geometrical representation of knowledge [27]. The principal motivation for introducing this layer (Fig.1) is the different abstraction levels between sub-symbolic and symbolic layers. The conceptual layer is an adaptation layer that facilitates the transformation of raw input data (text) into symbolic objects (operational concepts) for composing new concepts and discovering similarities

² nlp.stanford.edu/software

³ opennlp.apache.org

(concepts formation). We exploit this characteristic to fuse similar operational concepts and merge ontologies to construct the operational KG.

Concepts Formation

Concept formation is the construction of the ontology (the data graph schema) in a bottom-up approach that extracts knowledge instances from input textual data (ontology learning) and uses other knowledge resources coming from the CMs [28]. The whole KG is then created by ingesting data to the fused ontologies [29–31]. KGs are more abstract than CMs and allow unlimited connection of concepts in a flexible dynamic data structure. The interplay between CMs and KGs lead to concepts formation by merging the knowledge representation for CMs and the knowledge representation for KG:

- **Concept Maps** or conceptual diagram is a well-recognized approach [7] that uses both content knowledge and process knowledge to prompt users to create visual maps of a diagnostic strategy for identifying technical problems in complex environments [6]. A CM is a special type of propositional semantic network that is flexible, improves learning achievements, prompts constructive learning and active interactions [32, 33]. It is designed in the form of a directed graph where nodes represent concepts and edges represent relationships [34]. Despite being an old tool (1984), a CM visualizes the level of understanding and the level of thinking to assess the learning progress among human experts, displaying the structural nature and extent of knowledge, including misunderstandings of knowledge [32]. This aspect is quite important in our process because it facilitates human interactions and group synergies. The CM [7, 35] is the first step for the automatic generation of concepts from texts to *operational knowledge graph* where we follow the CM mining frameworks described in [34, 36]. As an example, in Fig.2 we present a fictitious example of a CM to illustrate the simple case of troubleshooting a not-lighting lamp.
- **Knowledge Graphs** are more recent (initially proposed by Google in 2012) and more adapted for automatic processing and machine reasoning than CMs. A KG is a knowledge base on a graph representing a set of concepts connected by relationships. The nodes are the entities or the concepts, while the edges represent the relationships among concepts for better knowledge organization [37]. A KG acquires and integrates information into an ontology and applies a reasoner to derive new knowledge [8]. Thus, there is no need to store all possible operational concepts and relationships in the KG because the logical reasoner can infer missing relationships at query time. KGs can efficiently express N-ary relationships between heterogeneous data in multiple domains using a hypergraph structure and clear denotation of entities, relations, and attributes. For example, in Fig.3, we present the KG schema for the use case of troubleshooting the not-lighting lamp. The main difference with the equivalent CM in Fig.2 is the new entity **SME** and the nested relationships for the abstract roles played (safety and operational) introduced to capture the need for SME supervision.

Fig. 2 Concept Map for the not-lighting lamp troubleshooting (*ContextMinds*).

Fig. 3 Knowledge graph schema for the not-lighting lamp troubleshooting (*TypeDB* from Vaticle Labs).

- **Ontology** promotes knowledge sharing [38] and offer a common communication mode for the knowledge elicitation process. The ontology describes and captures the domain knowledge [39], establishing the definitions of the technical concepts used by SMEs and providing the meaning of the relationships between technical terms and operational concepts. As such, ontologies aim to make domain knowledge explicit, remove contradictions and ambiguities, separate domain knowledge from operational knowledge, enable machines to reason and learn, and facilitate knowledge sharing between machines and humans [40]. Foundational domain ontologies have been developed for industrial domains such as aviation, aerospace, construction, steel production, chemical engineering, product development, and many others. A detailed presentation of the current state of ontologies for Industry 4.0 and reviews for existing ontological frameworks and ontological standardization efforts can be found in [41]. However, the typical condition is that a full ontology

for a particular industrial domain does not exist or is available only partially. In this case, the domain ontology is initially developed from the CMs [28] with the foundational ontologies used as guidance for domain experts during the collaborative process (Section 4) taking advantage of the many ontology learning methods from texts for extracting ontologies with NLP techniques [42].

3.1.3 Symbolic Layer

Once concepts have been formed in the conceptual layer, the SMEs and the knowledge engineers improve the knowledge models generated, looking for new concepts and relationships in a collaborative knowledge construction process. The symbolic layer is where the concepts initially formed in the conceptual layer (and sometimes ill-formed) are validated to create a consistent KG. Here we are mainly referring to ontology alignment [43, 44] expressed for making possible the integration into a data graph schema. The implicit knowledge is also converted into explicit knowledge and encoded in an ontological structure during the collaborative knowledge process with experts (Section 4). Therefore, the consequential data ingestion follows once the data graph schema has been generated to complete the KG generation. Actually, we follow an iterative process of learning cycles, and we learn how to build an enhanced KG improving the capabilities to integrate new fragments of knowledge extracted from the conceptual layer.

Logical Reasoner

Partial information is naturally stored in a graph where the relationships between operational concepts are integrated with a domain ontology, and implicit knowledge is inferred by the logical reasoner. In this paper, we follow the KG life cycle described in [45] with three distinct reasoning phases. In particular, in the learning phase (see Fig.4), we reason for knowledge integration and knowledge discovery (Section 4), and we use the logical reasoner to derive new knowledge, add missing knowledge to identify conflicting information generated in the conceptual layer (3.1.2). We note how some aspects of tacit knowledge can be encoded in the KG model as logical rules. For example, in the non-lighting lamp problem (see Fig.2 and Fig.3), there is no direct relationship between incandescent lamp and filament condition (broken or intact), but it can be added to the data schema with a rule stating that when the filament for an incandescent lamp is broken, then the lamp does not light-up. Finally, reasoning for application service is related to the operational phase (Section 4), when knowledge is retrieved for knowledge base question answering (KBQA) services or for search services to provide on-site training to new hires. A detailed discussion of reasoning in KGs can be found in [45, 46].

3.1.4 Development Software

We have used an exploratory programming approach to support the solution proposal described in this paper. We use the **ContextMinds**⁴ tool to merge CM and KG. We use **Text2Onto** [47] for ontology learning from text because knowledge is modeled at a meta-level and can be easily translated into different target languages. We store the KG on the graph database **TypeDB**⁵ that provides a strongly typed knowledge representation system based on hypergraphs and enables modeling any type of complex network in an ontological schema using entities, relations, attributes, and rules. For reasoning on KGs, we have also explored **RNNLogic** [48] for its capability to train a reasoning predictor with logic rules. For NLP analysis, we use **Stanza** [49], a Python NLP library for syntactic analysis.

4 Knowledge Elicitation Process

The knowledge elicitation process consists of a set of methods to elicit the tacit knowledge of a domain expert [50] with a mix of algorithmic techniques and organizational best practices. In particular, we focus on problem-solving knowledge [51], which is about capturing the domain knowledge of workers in a specific industrial structure during the accomplishment of tasks, such as maintenance operations, troubleshooting, and reparations. In management contexts, many different techniques have been proposed [10]. Our proposal goes beyond these techniques and describes an elicitation process based on a cognitive framework that transforms heterogeneous textual inputs and domain ontologies into a KG for knowledge retrieval. We follow a *role game* paradigm, in which human (H) and virtual (V) participants with different skill levels, from experts (E) to apprentices (A), play together. Our configuration echoes the renowned Turing’s imitation game but with a significant difference. We implement a role-playing game where the objective is not to assess if the virtual assistant has acquired some human characteristics, but rather to assess the correctness and reliability of the knowledge transferred between human and virtual agents and facilitate as much as possible the translation of SMEs tacit knowledge into valuable explicit knowledge through socialization and externalization (Section 3). The knowledge creation framework allows the transfer of insightful knowledge from SMEs to virtual agents in an iterative learning process under the supervision of human experts. This point is particularly critical because only human agents can diagnose why and how the virtual assistant may or may not be successful at specific tasks.

For the sake of completeness, this type of translation based on the SECI model is not new and has been described by other researchers [52], while the role game as a simulation of the professional activity with the participation of several experts has been described in [10] but without the involvement of the virtual agents. The participation of human experts in the knowledge

⁴www.contextminds.com

⁵vaticle.com

elicitation process will also safeguard against possible collateral effects. In some instances, the tacit knowledge employed by SMEs is incorrect or dangerous and shall not be exposed to the virtual assistant during the learning phase. Incorrectly used equipment or incorrectly interpreted results or procedural shortcuts can impose risks on the product, the service quality and negatively impact employees' and consumers' safety. This aspect goes beyond the logical reasoning capability of the virtual assistant and a purely algorithmic approach. Therefore, the knowledge engineers —the game's referee, will double-check the practices used by the SMEs and determine which knowledgeable facts can be kept and stored in the KG and which others must be questioned and avoided because dangerous, unsafe, or illegal. In this way, we ensure that the human apprentices can eventually interact with the cognitive system in the operational phase without the risk of learning uncontrolled facts that may compromise their learning experience and future productivity in the organization.

4.1 Knowledge Elicitation Work-Breakdown

The overall process for capturing tacit knowledge can be broken down into four phases (see Fig.4). The objective of the first three phases is to capture tacit knowledge into a KG in a role-playing game similar to a video conference. The last phase is not scoped in this paper, but briefly described.

- **Phase I - Dry Run:** In this phase, two human experts (*HE*) generate reference CMs that are compared against the expert's CM generated on similar use cases. We will assess whether the experts can sustain a rich technical exchange in a simulated working environment that may look unfamiliar. The rationale is to select the most knowledgeable experts on their subject of expertise and ensure optimal collaboration with the other team members (the apprentice and the knowledge engineer).
- **Phase II - Init:** We start by replacing the less performing *HE* with a knowledge engineer (*KE*) that will coordinate the elicitation process. The *KE* will also impersonate the apprentice. Selecting the *HE* and the *KE* that fit better is critical for the rest of the process. There may be experts who know *how to do* but cannot explain the process they do. To mitigate this possibility, we introduce the role of the *game master* played by the *KE*, who facilitates the communication exchange with SMEs. The *KE* should be technically knowledgeable and possess a positive command attitude and directional authority for creating a friendly environment and making the knowledge transfer possible. The natural choice is to select the knowledge engineer from the pool of the best-performing SMEs, but other options are possible that strongly depend on the corporate culture and personnel availability. The *HE* and the *KE* will also be involved in the creation of gold standards for quality assessment made of CMs extracted by human annotators from a set of tests [36].
- **Phase III - Learning:** In the learning phase, the virtual expert learns concepts and relationships from the role game technical session played by SMEs

Fig. 4 Phases for the Knowledge Elicitation Process: *HE* is the Human Expert, *HA* the Human Apprentice, *KE* the Knowledge Engineer, *VE* the Virtual Expert, and *VA* Virtual Apprentice. *KE_q* and *VA_q* are the Knowledge Engineer and the Virtual Apprentice asking questions, while *HE_a* and *VE_a* are the Human Expert and the Virtual Expert giving answers, respectively. Phases I, II, and III capture tacit knowledge into a KG, while Phase IV is the operational setup-up when the human apprentice interacts with the virtual expert (the cognitive assistant).

and knowledge engineers. We use CMs mining techniques [36] and ontology learning methods to generate a KG. During the process, the human expert (*HE*) is challenged by the *KE* impersonating the apprentice. The initial KGs are enhanced via an iterative Q&A session during which the participants exchange technical information until the *HE* and the *KE* agree that a successful knowledge transfer has been completed. Typical examples are found in machine service, industrial troubleshooting, production processes control, etc.

The human/machine interaction is done on two channels: *channel 1* between the virtual apprentice (*VA*) and the *HE*, and *channel 2* between the virtual expert (*VE*) and the *KE* (Fig. 5). The rationale is to isolate direct interactions of the human agents, which are gradually replaced along the process by their virtual counterparts impersonated by *avatars* depending on the level of quality reached in the sessions. The quality of the concepts produced in this phase will be scored by measuring the semantic similarity between concepts generated by the apprentice and the expert [53]. In particular, *channel 1* isolates the interactions of the human expert giving answers

Fig. 5 Phase III - Learning. On channel 1, HE_a is the Human Expert giving answers, VA_q the virtual apprentice asking questions impersonated by its avatar on the screen masking the Knowledge Engineer asking questions KE_q . On channel 2, VE_a is the Virtual expert giving answers impersonated by its avatar on the screen masking the human expert HE_a accordingly. The rectangle with dashed border captures the configuration used in Phase IV (operations) where we eventually replace KE_q with the apprentice HA .

(HE_a) and the knowledge engineer asking questions (KE_q) through the virtual apprentice (avatar) asking questions (VA_q), and *channel 2* isolates the interactions of the KE_q and the HE_a through the virtual expert (avatar) giving answers (VE_a). This arrangement in which humans are represented as avatars in a virtual environment and where each human sees the other as an avatar on their screen has been described in [54] in the context of an autonomous system for achieving artificial general intelligence. However, in our arrangement, the objective is to activate the experts' minds and reveal their tacit and implicit creative thinking procedures in a role-playing game [10]. Accordingly, the cognitive system learns how to virtualize human agents for knowledge transfer. We also note that in the process described, some of the implicit cultural values of the organization —the tacit knowledge, are implicitly captured and encoded in the knowledge base schema.

- **Phase IV - Operations:** In the operational phase, the VE replaces the HE , and the HA replaces the KE . In this phase, the KG is integrated into a cognitive system to allow new hires to formulate a technical problem as a collaborative task to the virtual expert, elicit mental models, and analyze the results [6] for knowledge retrieval and knowledge visualization [55]. Research done on regular classrooms has shown that learning with KGs resulted in better performance by students [32]. We expect the same improvements for workers in the industry interacting with this cognitive system for on-site training services.

5 Ethical and Societal Implications

Ethical and societal concerns are inevitable for the type of cognitive system we have described. Organizations should be vigilant about the knowledge management procedures for transferring tacit knowledge to be fair and equitable for human participants in the process. Once successfully trained, the cognitive system will operate in an industrial environment to allow new hires for on-site training. All the aspects of knowledge management shall be considered: from knowledge creation to knowledge transfer, from knowledge sharing to knowledge governance [56]. Should the system operate with a conversational AI user interface, impact assessment for the creation and use of the interface [57] shall be conducted by an independent organization before deploying the system. More specifically, from a societal and ethical standpoint, we can demarcate three points of interest that broadly track, data, model, and impact:

1. **Data** relates to concerns around what data is used and how the data is collected. Regarding what data is used, we note above that we propose using technical documentation and internal reports rather than video and audio assessment. Notwithstanding this, our approach lends itself to others using this kind of data. This is problematic because the collection of *emotive* data (such as verbal and facial expressions) requires surveillance of staff over long periods. The ethical concern here is one of consent and the appropriateness of the potential use of emotive data.
2. **Model** relates to the conceptual and symbolic layer we have discussed above. Here ground assumptions are made, which may be deemed contentious given that behavior analysis is occurring. Concerns with bias can be raised regarding the exclusion of various types of unconscious behavior such as routed in variations in customs, language use — here, the danger of excluding certain sources of tacit knowledge is what is of concern.
3. **Impact** relates to how tacit knowledge is used and the readiness with which the techniques of assessing non-algorithmic factors such as unconscious, unexplained knowledge can be abused. In essence, the rendering explicit of that which is implicit can be used to monitor *subliminally* and possibly thereby manipulate staff, a concern the EU AI Regulation draft (2021) raises as a critical concern [58]. In sum, these ethical and societal considerations can be addressed through accountability, transparency, and good governance mechanisms, such as those proposed in [59].

6 Conclusions

We have described a solution proposal for capturing operational best practices of experienced workers (a.k.a. the tacit or tribal knowledge) in industrial domains for knowledge transfer. We use domain ontologies for Industry 4.0 and reasoning techniques to discover and integrate new facts into a KG. We describe a concepts formation iterative process as a role game played by subject matter experts and knowledge engineers interacting indirectly with a virtual

agent represented by an avatar. At the end of a learning phase, the expert is replaced by the virtual agent, and the role of knowledge engineer is impersonated by new hires or workers that need on-site training. Societal and ethical concerns have also been discussed.

7 Future Directions

We plan to consolidate the investigations performed and develop the complete cognitive architecture proposed in this paper. We also plan to extend the sensory input modalities to video stream sources for using T-Patterns analysis [60], adopting the approach described in [54] that may be used to improve the completeness of the KGs for troubleshooting tasks during the learning session with a better identification of temporal and sequential patterns —typical of manufacturing industrial processes.

Declarations

- Funding: No funding was provided for the specific research described in this article. The manuscript was made outside of work hours and using the authors' resources.
- Conflict of interest: The authors declare that there are no conflicts of interest or competing interests associated with this article or with the research it describes.
- Ethics approval: Not Applicable
- Consent to participate: Not Applicable
- Consent for publication: Not Applicable
- Availability of data and materials: Data sharing not applicable to this article as no datasets were generated or analysed during the current study.
- Code availability: Not Applicable
- Authors' contributions: EF conceived the initial cognitive framework and Section 4. HL contributed to Section 3. EK and AK are the authors of Section 5. The first draft of the manuscript and the following amendments was written by EF. All authors read and approved the final manuscript.

References

- [1] Santos, C., Mehraei, A., Barros, A.C., Araújo, M., Ares, E.: Towards industry 4.0: an overview of European strategic roadmaps. *Procedia Manufacturing* **13**, 972–979 (2017). Manufacturing Engineering Society International Conference 2017
- [2] Brockmann, E.N., Anthony, W.P.: Tacit knowledge and strategic decision making. *Group & Organization Management* **27**(4), 436–455 (2002)

- [3] Hadjimichael, D., Tsoukas, H.: Toward a better understanding of tacit knowledge in organizations: Taking stock and moving forward. *Academy of Management Annals* **13**(2), 672–703 (2019)
- [4] Polanyi, M., Sen, A.: *The Tacit Dimension*. University of Chicago Press, Chicago, IL (2009)
- [5] Nonaka, I.: A dynamic theory of organizational knowledge creation. *Organization Science* **5**(1), 14–37 (1994)
- [6] Yuetong Lin, Shahhosseini, A.M., Badar, M.A., Foster, T., Dean, J.: A concept map-based cognitive framework for acquiring expert knowledge in industrial environment. In: 2016 IEEE Frontiers in Education Conference (FIE), pp. 1–5 (2016)
- [7] Novak, J.D., Gowin, D.B., Kahle, J.B.: *Learning About Learning*, pp. 1–14. Cambridge University Press, UK (1984)
- [8] Ehrlinger, L., Wöß, W.: Towards a definition of knowledge graphs. In: Martin, M., Cuquet, M., Folmer, E. (eds.) *SEMANTiCS 2016*, Leipzig, Germany, September 12–15, 2016. CEUR Workshop Proceedings, vol. 1695 (2016)
- [9] Paulheim, H.: Knowledge graph refinement: A survey of approaches and evaluation methods. *Semantic Web* **8**, 489–508 (2017)
- [10] Gavrilova, T., Andreeva, T.: Knowledge elicitation techniques in a knowledge management context. *J. Knowl. Manag.* **16**(4), 523–537 (2012)
- [11] Dorobăţ, I.: *The role of tacit knowledge management in ERP systems implementation*. IO: Productivity (2008)
- [12] Barthelmey, A., Störkle, D., Kuhlenkötter, B., Deuse, J.: Cyber physical systems for life cycle continuous technical documentation of manufacturing facilities. *Procedia CIRP* **17**, 207–211 (2014)
- [13] Sanyal, C., Rigby, C.: E-mentoring as a HRD intervention: an exploratory action research study within an international professional mentoring scheme. *Human Resource Development International* **20**(1), 18–36 (2017)
- [14] Chen, P., Lu, Y., Zheng, V.W., Chen, X., Yang, B.: KnowEdu: A system to construct knowledge graph for education. *IEEE Access* **6**, 31553–31563 (2018). <https://doi.org/10.1109/ACCESS.2018.2839607>
- [15] Baumard, P.: Tacit knowledge in professional firms: the teachings of firms in very puzzling situations. *J. Knowl. Manag.* **6**(2), 135–151 (2002)
- [16] Gherardi, S., Nicolini, D.: To transfer is to transform: The circulation of

- safety knowledge. *Organization* **7**(2), 329–348 (2000)
- [17] Pozzali, A.: Tacit knowledge, implicit learning and scientific reasoning. *Mind Society* **7**(2), 227–237 (2007)
 - [18] Viale, R.: Tacit Knowledges, pp. 305–323. Springer, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40216-6_12
 - [19] Nonaka, I., Takeuchi, H.: *The Knowledge-creating Company: How Japanese Companies Create the Dynamics of Innovation*, p. 284. Oxford University Press, New York (1995)
 - [20] Nonaka, I., von Krogh, G.: Tacit knowledge and knowledge conversion: Controversy and advancement in organizational knowledge creation theory. *Organization Science* **20**(3), 635–652 (2009)
 - [21] Zárraga, C., García-Falcón, J.M.: Factors favoring knowledge management in work teams. *J. Knowl. Manag.* **7**(2), 81–96 (2003)
 - [22] Sakiroglu, M., Riedel, J., Pawar, K.S.: The knowledge creation process and model within NPD teams. In: 2005 IEEE International Technology Management Conference (ICE), pp. 1–10 (2005)
 - [23] Gourlay, S.: Conceptualizing knowledge creation: A critique of Nonaka’s theory. *Journal of Management Studies* **43**(7), 1415–1436 (2006)
 - [24] Fenstermacher, K.D.: The tyranny of tacit knowledge: What artificial intelligence tells us about knowledge representation. In: *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, pp. 243–243 (2005). <https://doi.org/10.1109/HICSS.2005.620>
 - [25] Susskind, D.: *Re-thinking the capabilities of machines in economics*. Department of Economics Discussion Paper Series. University of Oxford (2017)
 - [26] d’Avila Garcez, A.S., Besold, T.R., Raedt, L.D., Földiák, P., Hitzler, P., Icard, T., Kühnberger, K., Lamb, L.C., Miikkulainen, R., Silver, D.L.: Neural-symbolic learning and reasoning: Contributions and challenges. In: 2015 AAAI Spring Symposia, Stanford University, Palo Alto, California, USA, March 22–25, 2015 (2015)
 - [27] Gardenfors, P.: Conceptual spaces as a framework for knowledge representation. *Mind and Matter* **2**(2), 9–27 (2004)
 - [28] Soares, A., Sousa, C.: Using concept maps for ontology development: A case in the work organization domain. In: Azevedo, A. (ed.) *Innovation in Manufacturing Networks*, pp. 177–186. Springer, Boston, MA (2008)

- [29] Buchgeher, G., Gabauer, D., Gil, J.M., Ehrlinger, L.: Knowledge Graphs in Manufacturing and Production: a Systematic Literature Review. *IEEE Access* **9**, 55537–55554 (2021)
- [30] Zhao, Z., Han, S.-K., So, I.-M.: Architecture of knowledge graph construction techniques. *International Journal of Pure and Applied Mathematics* **118**(19), 1869–1883 (2018)
- [31] Zhao, M., Wang, H.-Q., Guo, J., Liu, D., Xie, C., Liu, Q., Cheng, Z.: Construction of an industrial knowledge graph for unstructured chinese text learning. *Applied Sciences* (2019)
- [32] Cui, J., Yu, S.: Fostering deeper learning in a flipped classroom: Effects of knowledge graphs versus concept maps. *British Journal of Educational Technology* **50**(5), 2308–2328 (2019)
- [33] Daley, B.J.: Concept maps: Practice applications in adult education and human resource development. *New Horizons in Adult Education and Human Resource Development* **24**(2-4), 31–37 (2010)
- [34] Zubrinic, K., Kalpic, D., Milicevic, M.: The automatic creation of concept maps from documents written using morphologically rich languages. *Expert Syst. Appl.* **39**(16), 12709–12718 (2012)
- [35] Novak, J.D., Cañas, A.J.: The origins of the concept mapping tool and the continuing evolution of the tool. *Inf. Vis.* **5**(3), 175–184 (2006)
- [36] Villalon, J.J., Calvo, R.A.: Concept map mining: A definition and a framework for its evaluation. In: 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, vol. 3, pp. 357–360 (2008)
- [37] Hogan, A., Blomqvist, E., Cochez, M., d’Amato, C., de Melo, G., Gutiérrez, C., Gayo, J.E.L., Kirrane, S., Neumaier, S., Polleres, A., Navigli, R., Ngomo, A.N., Rashid, S.M., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., Zimmermann, A.: Knowledge graphs. *Communications of the ACM* **64**, 96–104 (2021)
- [38] Hoppe, T., Eisenmann, H., Viehl, A., Bringmann, O.: Shifting from data handling to knowledge engineering in aerospace industry. In: 2017 IEEE International Systems Engineering Symposium (ISSE), pp. 1–6 (2017). <https://doi.org/10.1109/SysEng.2017.8088312>
- [39] Lian-dong, Z., Qifeng, W.: Knowledge discovery and modeling approach for manufacturing enterprises. 2009 Third International Symposium on Intelligent Information Technology Application **1**, 291–294 (2009)

- [40] Persson, C., Wallin, E.O.: Engineering and business implications of ontologies — a proposal for a minimum viable ontology. In: 2017 13th IEEE Conference on Automation Science and Engineering (CASE), pp. 864–869 (2017). <https://doi.org/10.1109/COASE.2017.8256212>
- [41] Sampath Kumar, V.R., Khamis, A., Fiorini, S., Carbonera, J.L., Olivares Alarcos, A., Habib, M., Goncalves, P., Li, H., Olszewska, J.I.: Ontologies for industry 4.0. *The Knowledge Engineering Review* **34**, 17 (2019)
- [42] Konys, A.: Knowledge repository of ontology learning tools from text. *Procedia Computer Science* **159**, 1614–1628 (2019). <https://doi.org/10.1016/j.procs.2019.09.332>. Knowledge-Based and Intelligent Information Engineering Systems: Proceedings of the 23rd International Conference KES2019
- [43] Husáková, M., Bureš, V.: Formal ontologies in information systems development: A systematic review. *Information* **11**(2) (2020). <https://doi.org/10.3390/info11020066>
- [44] Ngo, D., Bellahsene, Z.: Efficient semantic verification of ontology alignment. In: 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), vol. 1, pp. 141–148 (2015). <https://doi.org/10.1109/WI-IAT.2015.92>
- [45] Bellomarini, L., Sallinger, E., Vahdati, S.: Chapter 6 Reasoning in Knowledge Graphs: An Embeddings Spotlight, pp. 87–101. Springer, Cham (2020)
- [46] Chen, W., Xiong, W., Yan, X., Wang, W.Y.: Variational knowledge graph reasoning. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1, pp. 1823–1832 (2018)
- [47] Cimiano, P., Völker, J.: Text2onto. In: Montoyo, A., Muñoz, R., Métails, E. (eds.) *Natural Language Processing and Information Systems*, pp. 227–238. Springer, Berlin, Heidelberg (2005)
- [48] Qu, M., Chen, J., Xhonneux, L.-P., Bengio, Y., Tang, J.: Rnnlogic: Learning logic rules for reasoning on knowledge graphs. In: International Conference on Learning Representations (2021)
- [49] Qi, P., Zhang, Y., Zhang, Y., Bolton, J., Manning, C.D.: Stanza: A Python natural language processing toolkit for many human languages. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations (2020)

- [50] Shadbolt, N., Smart, P.R.: In: Wilson, J.R., Sharples, S. (eds.) *Knowledge Elicitation: Methods, Tools and Techniques*, pp. 163–200. CRC Press, ??? (2015)
- [51] Sureephong, P., Ouzrout, Y., Buzon, L., Bouras, A.: Knowledge management system enabling collaboration in industry cluster. *Int. J. Manuf. Res.* **5**(4), 478–497 (2010)
- [52] Kimmerle, J., Cress, U., Held, C.: The interplay between individual and collective knowledge: technologies for organisational learning and knowledge building. *Knowledge Management Research Practice* **8**(1), 33–44 (2010)
- [53] Zhu, G., Iglesias, C.: Computing semantic similarity of concepts in knowledge graphs. *IEEE Transactions on Knowledge and Data Engineering* **29**, 72–85 (2017)
- [54] Nivel, E., Thórisson, K.R., Steunebrink, B.R., Dindo, H., Pezzulo, G., Rodríguez, M., Hernández, C., Ognibene, D., Schmidhuber, J., Sanz, R., Helgason, H.P., Chella, A., Jonsson, G.K.: Bounded recursive self-improvement. *CoRR* **abs/1312.6764** (2013) [arXiv:1312.6764](https://arxiv.org/abs/1312.6764)
- [55] Sun, K., Liu, Y., Guo, Z., Wang, C.: Visualization for knowledge graph based on education data. *International Journal of Software and Informatics* **10**(3) (2016)
- [56] Rhem, A.J.: AI ethics and its impact on knowledge management. *AI Ethics* **1**(4), 33–37 (2021)
- [57] Ruane, E., Birhane, A., Ventresque, A.: Conversational AI: Social and ethical considerations. *AICS* (2019)
- [58] European Commission: Proposal for a Regulation on a European approach for Artificial Intelligence (2021). <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-approach-artificial-intelligence>
- [59] Koshiyama, A.S., Kazim, E., Treleaven, P.C., Rai, P., Szpruch, L., Pavey, G., Ahamat, G., Leutner, F., Goebel, R., Knight, A., Adams, J., Hitrova, C., Barnett, J., Nachev, P., Barber, D., Chamorro-Premuzic, T., Klemmer, K., Gregorovic, M., Khan, S.A., Lomas, E.: Towards algorithm auditing: A survey on managing legal, ethical and technological risks of AI, ML and associated algorithms. *Software Engineering eJournal* (2021)
- [60] Magnusson, M.S.: T-Pattern detection and analysis (TPA) with THEMETM: A mixed methods approach. *Frontiers in Psychology* **10**, 2663 (2020). <https://doi.org/10.3389/fpsyg.2019.02663>