How much Knowledge is in Knowledge Graphs? - A Knowledge Management Perspective

Rositsa V Ivanova *, Clemens Kerschbaum, Sebastian Neumaier, Alexander Kaiser, and Axel Polleres
Vienna University of Economics and Business, Welthandelsplatz 1, Vienna, Austria
E-mails: rositsa.ivanova@wu.ac.at, clemens.kerschbaum@wu.ac.at, sebastian.neumaier@wu.ac.at, alexander.kaiser@wu.ac.at, axel.polleres@wu.ac.at

Abstract. Managing and preserving knowledge in the best possible way has always been a key to the success for organisations, long before the term “Knowledge Graph” has entered the stage. However, the understanding of what exactly knowledge is, how it is represented and organised, and how knowledge is created often varies between different research communities. To this day, the scientific discipline of Knowledge Management is trying to capture the process of knowledge creation as converting implicit, i.e., tacit, knowledge into explicit knowledge. In this paper, we first give an idea of this Knowledge Management perspective on knowledge creation, and then discuss how Knowledge Graphs actually can contribute to solve the issue of making implicit knowledge explicit. We empirically survey the use of Knowledge Graphs in enterprise environments and – picking three concrete examples – discuss concrete use cases from a Knowledge Management viewpoint.

Keywords: Knowledge Graph, Knowledge Management, Knowledge Creation

1. Introduction

The term and concept of a “Knowledge Graph” (KG) has recently been happily embraced by the Semantic Web [1] research community. It has opened new directions beyond Web standards for semantic links between URI-identified concepts alone (i.e., Linked Data), towards organizing data around semantic concepts and their relationships in more general terms. This said, we note that other also core computer science sub-communities have either used the term “knowledge” much longer already, or recently start to adopt it.

The Knowledge Representation and Reasoning (KRR) community, for example, works under the “fundamental assumption [...] that an agent’s knowledge is explicitly represented in a declarative form, suitable for processing by dedicated reasoning engines.”¹ Yet, the “Knowledge Graph” metaphor seems to be broader than the KRR understanding, also spilling over to other communities who recently pick up the term. This is for instance demonstrated with articles carrying the term “Knowledge Graph” in their titles in top conferences on the topics of Databases (from this year’s SIGMOD and VLDB conferences) [2–4], Machine Learning (from this year’s NIPS conference) [5–7], or natural language processing (from this year’s ACL conference) [8–10].

These references deliberately only reflect a current snapshot of communities and approaches, that may not have focused on considerations of explicit knowledge in the past, yet now recognise knowledge itself as something that could be “created” or “completed” by automated means. Such consideration of explicit

Factual knowledge in a KG is often approached by embeddings, that is, representations of KGs in a vector space, amenable to sub-symbolic processing using (deep) neural networks.

Thus one could argue that while the core KRR community understands knowledge as something that is already explicit, formalised, and that can be operationalised by automated agents, this new KG view is also working with sub-symbolic, rather – as one could claim – implicit forms of knowledge. In particular, the core novelty seems to be combination of explicit and implicit knowledge, which is potentially leveraged through KGs. This new perspective has fueled the hope for an artificial but still explainable intelligence in machines [11], where such “explainability” should be understood as again being explicitly communicable to humans.

Still, we argue that many research works in the area of and around KGs do not really define what they mean by “knowledge” itself. We believe that such a definition is the very basis for assessing to what extent “knowledge” can indeed be created by machines. Apart from computer science, fortunately, the field of Knowledge Management (KM), rooted in organisational sciences, has dealt with exactly this question since decades already.

Looking at the history of KM, the knowledge-based view on organisations builds on the foundations of the resource-based theory of the firm, introduced by Penrose [12] in 1959. Penrose states that knowledge is crucial for every firm because it enables the use of all tangible resources in order to create value. The better an organisation instrumentalises given resources through knowledge, the more likely it is to acquire competitive advantages in the long run [13–15]. However, the problem is that knowledge in a firm is hard to grasp – there is nothing like a “catalogue of knowledge”. In fact, most knowledge is invisible as it is often embedded in organisational routines, carried out as organisational culture or simply bound to individual employees [16]. Usually, in relation to all existing knowledge in a company, only a very small amount of knowledge is codified and made explicit. In other words “we can know more than we can tell”, as Polanyi so thoroughly put it [17], which implies that the share of implicit knowledge in organisations is much higher than the share of explicit knowledge and that tacit knowledge is a source of competitive advantage for organisations [18]. Thus it can turn out difficult to access a firms knowledge because we firstly do not know what is all there, secondly if we knew what was there, we would not know where it was and thirdly it would eventually be hard to make sense out of it without further knowledge of the context around a specific chunk of knowledge.

In order to tackle all those problems, the KM literature since the turn of the millennium and onward points out that organisations should strategically manage their knowledge [16].

It also showed that technology proves as valuable means to exploit knowledge [19]. Especially in terms of storing already codified knowledge but also to connect individuals in an organisation with experts in order to leverage their tacit knowledge on a certain subject [16, 20]. Recently social media had also been taken into consideration when it comes to locating knowledge as an answer to the “who knows what” question [21].

However, as Alavi and Leidner point out “it is less the knowledge existing at any given time per se than the firms’ ability to […] create new knowledge […]” that lead to competitive advantage from knowledge-based assets [16]: it is the generation of new knowledge that really creates value. A widely appreciated explanation for the process of knowledge creation is Nonaka’s SECI model [22]. Following this model, the evolution of knowledge is a circular process involving the recurring transformation of tacit knowledge into explicit knowledge, and explicit knowledge back into tacit knowledge that leads to the creation of new knowledge [22]. The process of knowledge transformation along the SECI model particularly emphasises the social component of knowledge exchange, rather than technical means.

Although there is a lot of literature in the KM community about how tacit knowledge can be used in organisations and transformed into explicit knowledge [23], the question whether modern technical means – in particular KGs – are suitable and supportive for this task still remains largely unanswered.

In order to close this gap, in the present work we focus on the question:

How can knowledge graphs foster the conversion of implicit into explicit knowledge and thus support the generation of new knowledge in organisations?

We herein aim at examining how KGs could become the key to make an organisation’s knowledge explicit and how they can provide us with possibilities to iden-
2. Knowledge Management

Discussions on what is “knowledge” are getting more relevance than ever due to the aforementioned trending technology areas, whereas other disciplines have a long before dealt with this question from a much broader perspective. In particular, the socio-technical perspective of the concerned domains is often overlooked in recent discussions: while such a perspective is often regarded as a rather philosophical one, there is pragmatic relevance to it. It discusses questions such as “What does a robot need to know in order to open a safe, and how does it know whether it knows enough to open it?” [24]. The answers to such questions include the knowledge of multiple actors within groups. Therefore, we will further regard the term knowledge both on a personal and on an organisational level.

An organisation is typically regarded as a group of people. On an organisational level it holds knowledge, consisting of the constantly repeated routines of the organisation - called “organisational knowledge” [25]. On an individual level, all members of the group are holders of individual knowledge, contributing to the creation of new and the forwarding of existing knowledge. At the same time, the inarguably important role of automated means and information systems, which continuously provide people with new information in this knowledge creation process needs to be considered. In this paper we take a step beyond the boundaries of an organisation and discuss the effect of the surrounding environment and it’s stakeholders.

2.1. What is Knowledge?

The term “knowledge” and its definition date back to the classical philosophers of the Greek era and are still actively discussed in recent publications. In Theaetetus by Plato (369 BC), Socrates introduced the traditional definition of knowledge as “justified true belief” (i.e., subjective personal perspective) [26, 27]. Rather than a simple compilation of facts, the creation of knowledge is characterized by an individual making sense of a new situation by holding a true belief with an account [28]. Later on, this understanding is challenged by the renowned publication of Edmund Gettier [29]. The philosopher presents two counterexamples, illustrating that even justified true belief could turn out to be false. Consequently, a cognitive perspective would argue that the knowledge is dependent on the individual making sense of it [30].

Another definition of knowledge views it as an enabling activity, a “capacity to act” [31]. Sveiby focuses on the action element and thus on the dynamic properties of knowledge. He argues that knowledge is only useful, when it is available to individuals and increases their capacity to act. All those and further perspectives on knowledge have lead to multiple definitions on individual and organisational levels.

Data, information, knowledge The literature on knowledge management frequently distinguishes between data, information, and knowledge. Faucher et al. give an overview of the definitions of those terms [26]. While there does not seem to be a consensus, similarities in the understanding can be found.

- Data is primarily seen as unprocessed raw representation of reality.
- Information builds on it by providing a context and thus a semantic meaning.
- Once the gained information is utilised pragmatically for the introduction of new actions, new meanings, etc. we speak of knowledge.

Faucher et al. also discuss the term wisdom in detail; we excluded the concept of wisdom in our discussions as it lies out of the scope of the paper.

Knowledge is carried through multiple entities (e.g. organisational routines, culture), is context specific and dependent on a particular time and space [27, 32]. Without those characteristics, it simply becomes information. Various publications in the past have stressed upon the importance of differentiating between knowledge and data or information. In contrast to information, knowledge is about beliefs and commitment, as well as about action and lastly, about meaning [33]. If knowledge is not different from data or information, it could be described as a stock rather than a flow and an opportunity for the generation of something new or
interesting in organisational terms [34]. Additionally, there are two aspects of information - the syntactic and the semantic one. The syntactic aspect merely observes the volume of the information flow, while the semantic aspect conveys a meaning, which is useful for the creation of knowledge [33].

Implicit and Explicit Knowledge Taking a different viewpoint, the literature on knowledge management has agreed upon the separation of knowledge in two yet interconnected concepts - namely, explicit knowledge and tacit knowledge [35]. On the one hand, explicit knowledge can be expressed by individuals in formal and systematic language in forms such as text and sound, and is thus easily communicated, stored and processed. On the other hand, tacit knowledge rests in individuals’ intuitions, actions, bodily experience and also in organisations’ routines and habits. In contrast to explicit knowledge, tacit knowledge is not easily made expressible [27, 33]. As pointed out by Virtanen [36], one may find multiple definitions of tacit and explicit knowledge throughout the KM literature. In this paper we adopt the definition from Polanyi’s theory without discussing the philosophical foundation it builds upon. It describes those two types of knowledge as different kinds of awareness - a continuous spectrum - rather than as separate or even independent concepts. In this context, explicit knowledge refers to focal awareness, tacit - to subsidiary awareness. Those two are mutually exclusive, this means that the awareness cannot be both simultaneously. Nevertheless, when the attention is put on the tacit knowledge, it turns from subsidiary to focal (i.e. from implicit to explicit). Lastly, subsidiary awareness serves as an enabler of focal awareness through its various clues, elements and processes [36]. In order to better understand the difference between explicit and tacit knowledge, it is beneficial to outline how new knowledge is created.

2.2. On Knowledge Creation

The creation of organisational knowledge can be separated into two important steps - the process of knowledge creation by individuals and the integration of individual knowledge into that of an organisation [18]. The famous SECI model (Socialisation, Externalisation, Combination, Internalisation) by Nonaka and Takeuchi identifies four steps of this knowledge creation process [35]. From a two-dimensional perspective an organisation converts tacit knowledge into explicit knowledge and then back to tacit knowledge. In the Socialisation mode tacit knowledge is shared and created by individuals through their actions and observations. Next, dialogue, reflection, metaphors are used to make tacit knowledge explicit. This is referred to as Externalisation. In the third mode - the Combination - explicit knowledge is structured and applied. Lastly, the explicit knowledge shifts to tacit by being simulated, applied and thus embodied in the Internalisation phase [27, 33].

In addition to the those two dimensions of the SECI model, the process contains an ontological level, which turns it into a spiral. Beyond the differentiation between explicit and tacit knowledge, each phase of the model captures the interaction between different combinations of entities - individual, group and organisation. While the Socialisation phase takes place only between individuals, each of the remaining phases involves a different pair combination of the three entities (in the given order) [27, 33].

2.3. Technical Support in the Process

The management of existing and the creation of new knowledge is often supported on different levels by information systems. In their work on how organisations manage knowledge, Davenport and Prusak touch upon the ways in which computers may help “transform data into information by adding value to it” [37]. They recognise a total of five relevant methods – Contextualized, Corrected, Categorized, Calculated, and Condensed – noting that the latter three methods must be performed by humans, as only an individual could add meaning to data [37].

Considering the progress of technology over the last decades, however, it may become arguable, whether there is still a necessity for frequent human intervention in any of those above-mentioned five methods. In addition, it may be arguable that information systems merely contain information (only). However, it is important to consider that most so-called knowledge bases (as a term in between “database” and “knowledge”), in particular most KGs, have often been created by individuals (in companies or crowdsourced) and thus store their collective knowledge. The stored knowledge is then made use of by organisations for their individual purposes. In the following subsection we introduce and discuss Knowledge Graphs as a recently emerging approach for storing, structuring and organising knowledge bases in a graph in more detail.
3. Knowledge Graphs

While indeed the term "Knowledge Graph" appears in the literature already for decades (at least since 1974 [38]), however, the recent uptake of the term stems from industry: Google’s announcement of their Google Knowledge Graph in 2012 initiated several subsequent announcements and developments of KGs, e.g., at Airbnb, Amazon, Facebook, Uber, etc. [39]. This uptake in industry was followed by many scientific publications, including articles on definitions of KGs [40], on construction and refinement techniques [41], books [42–44], etc. We try to give an idea and understanding of what is covered by the term, by first discussing some definition approaches and then analyzing the underlying principles and techniques of modeling and creating KGs.

Definitions and Graph Models The idea behind the concept of KGs – both in the enterprise setting and in academia – is to represent information in a graph abstraction. Existing graph models offer a number of benefits compared to relational models and NoSQL alternatives. Hogan et al. propose the following – very inclusive – view on KGs: “a graph of data intended to accumulate and convey knowledge of the real world, whose nodes represent entities of interest and whose edges represent relations between these entities.” [39]

A more concrete definition is given by Ehrlinger and Wöß by defining a KG as a graph that “acquires and integrates information into an ontology and applies a reasoner to derive new knowledge.” [40] This definition distinguishes a KG from an ontology by the ability of generating new knowledge through reasoning.

There are different graph-based data models for building a KG, such as directed edge-labelled graphs, and property graphs. A directed edge-labelled graph is defined as a set of nodes and a set of directed edges, where both nodes and edges are labelled. Typically, nodes represent entities and the edges relations between them. A standardised model for directed edge-labelled graphs is RDF [45] and the corresponding query language SPARQL [46]. A prominent application of RDF is the Linked Open Data Cloud,3 which consists of various interlinked, queryable, RDF documents on the Web.

Property graphs are an alternative, generalised graph model which were introduced to provide flexibility in terms of additional information associated with the edges and nodes [47], i.e., edges and nodes have additional property-value pairs, often used to convey for instance contextual or provenance information, attached. This allows a more compact and flexible modelling in certain cases, however, directed edge-labelled provide a more minimal – in principle equally expressive – model [39]. The property graph model is particularly popular among some widely used graph databases such as Neo4j.

Creation of Knowledge Graphs The techniques to create and enrich KGs range from automated methods, to human collaboration and crowd-sourcing. The methodology heavily depends on factors such as the domain of data, the application of the KG, and the involved actors, or, resp., the availability of underlying data sources. Some popular existing KGs have been constructed from human collaboration. For instance, the Wikidata project4 is a graph of manually curated and updated entities and relations across different domains and languages. Human collaborators collect and edit the information in Wikidata. There is an interface to issue structured queries; the information from Wikidata is then re-used and linked in various other sources on the Web, e.g., in Wikipedia articles.

Complementary, the popular DBpedia project [48] is based on an automated extraction of the structured information in Wikipedia. The extraction constructs a graph based on the links (i.e. relations) from one article (i.e. entity) to another within Wikipedia, leveraging the template structure of infoboxes to label relations. The resulting KG is further automatically enriched by linking to further structured resources from e.g. GeoNames5 and WordNet [49].

Alternative KG creation techniques are based on entity linking and relation extraction from text sources, as well as mapping and relation extraction from Web tables, etc. [39].

For a comprehensive survey covering various aspects including the construction, models, ontology and entailment techniques we refer to the survey by Hogan et al. [39]. For further reading on KG refinement techniques we refer to the survey by Paulheim [50]; regarding applications and use cases we refer for instance to the books by Fensel et al. [44] or Janev et al. [43].

---

3https://lod-cloud.net/
4https://www.wikidata.org/wiki/Wikidata:Main_Page
5https://www.geonames.org/
4. Management and Creation of Knowledge using Knowledge Graphs

Apart from use cases on the (public) Web, KGs are increasingly also used within companies and organisations to structure data and manage their knowledge. In order to explore how organisations can leverage KGs for actual knowledge creation in the field, we therefore surveyed conference publications published by companies. The goal of this survey is to identify how KGs help to manage organisational knowledge, and eventually highlight cases where KGs help to indeed make implicit knowledge explicit and thereby generate new knowledge.

4.1. Literature Survey

In order to assess the development of applications of KG applications, we focused on relevant publications in the International Semantic Web Conference (ISWC) as a top ranked conference in the field of semantic web and one of the main venues for publications related to KGs. Particularly, since ISWC also provides a dedicated industry track for presenting solutions based on and adoption of semantic technologies in applied settings, we focused along the trending topic of KGs on the increasing number of papers that describe the application of KGs in specific industry domains.

For our survey, we considered all available publications of the ISWC industry track for the years 2019, 2018, and 2017 online. In total, we surveyed 48 papers, as shown in Table 1 in order to obtain insights in how far KG technologies reportedly contribute to knowledge creation within organizations.

<table>
<thead>
<tr>
<th>year</th>
<th>papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019</td>
<td>16</td>
</tr>
<tr>
<td>2018</td>
<td>14</td>
</tr>
<tr>
<td>2017</td>
<td>18</td>
</tr>
<tr>
<td>total</td>
<td>48</td>
</tr>
</tbody>
</table>

Table 1
Number of papers surveyed per year.

Categorisation & Metrics

For each of the papers we collected a set of – in our opinion – relevant features and characteristics. The list of all publications and our respective assessment can be found online.

- **term KG:** Firstly, we scanned the papers for the explicit use of the term “knowledge graph”. Our rationale behind this is to highlight these papers, however, we also observed relevant publications using a different terminology.

- **data sources:** Secondly, we categorise the source of the described/covered data into internal data, external data, or both. We refer to internal data as data previously existing or produced within the company. In contrast, external data is data that has not previously been integrated in the company’s work. We distinguish between these cases in order to highlight whether internal or external sources were used for creating the KG and thus for the potential creation of new knowledge.

- **open KGs:** In case the paper explicitly describes the integration/use of external sources, we further consider if it is explicitly stated that the external data comes from an open KG (such as DBpedia or Wikidata).

- **knowledge creation:** We are particularly interested in papers that describe a process of making knowledge explicit through the use of KG techniques.

In Table 2 we quantitatively summarize the surveyed ISWC industry track papers. In 20 papers (out of a total of 48) the term “knowledge graph” was used. This indicates a strong focus on the use of KGs in industry in recent years at ISWC. Most of the reviewed publications discuss the use/integration/representation of data from internal sources. We only observed 12 papers (in the external and both category) which mentioned the use/integration non-internal data sources; eight out of these papers use openly available KGs (e.g. Wikidata, or DBpedia). Of particular interest to our research were the 16 papers (category knowledge creation) that explicitly discuss how KGs helped to make knowledge explicit in an enterprise context.

---

10Note: The sum of the data sources in the table (i.e., internal, external and both) does not correspond to the number of papers in the years 2017 and 2019, since in some of the papers, this categorisation did not appear to be applicable or it turned out there was no explicit description of the used data sources.
In addition to the selected metrics, we collected metadata from the surveyed papers such as the keywords provided by the authors, the company names for the corresponding paper and the research area of the published paper. While those metrics are not directly needed for the analysis/diagnosis in this paper, we find them helpful for the purpose of getting a better overview of the applications of KGs in industry. More details can be found in the spreadsheet provided online.

4.2. Case Studies

To exemplify the process of “knowledge creation” – i.e., gaining new explicit knowledge – using the idea and techniques of KGs, we discuss a selection of three of the surveyed papers as a small, but in our opinion representative case studies: (i) In Strötgen et al. [51] the sources of knowledge, i.e. relational databases and unstructured documents, get organized in a KG to generate new knowledge and eventually improve search functionalities. (ii) In Cotroneo et al. [52] the already structured information from several sources get integrated into the company’s KG; again to improve and increase the search scope. Finally, (iii) In Ireson and Ciravegna [53] existing open KGs are integrated: Wikidata and DBpedia are used to enrich the initial domain knowledge base and to more accurately solve a prediction task. We consider those three to be representative, in the sense that the other articles we surveyed resemble these – general complementary – three uses, i.e., (i) knowledge generation by re-structuring, (ii) knowledge generation by integration, and (iii) knowledge generation by enrichment.

(i) Strötgen et al., 2019. Towards the Bosch Materials Science Knowledge Base [51]

In the context of manufacturing, there is a wide variety of materials used in the creation of innovative products. Due to the rapid introduction of new papers, patents and regulations in the area, the retrieval of relevant information may be challenging. Often, complex queries are necessary to find the appropriate answers (e.g., “Find anode materials in Intermediate Temperature Solid Oxide Fuel Cells (IT-SOFC) that produce high power density” [51]). The authors state three required goals for such a query answering system: (i) integration of different sources into a unified KG, (ii) complex query functionalities such as aggregation and multi-hop reasoning, and (iii) provenance for query results.

Combination of Data Sources The paper describes the use of both internal and external sources that get integrated in a graph. The system integrates internal relational databases using Ontology-based access methods. Further, relevant text-based publications and patents are processed and integrated using NLP techniques.

Creation of new Knowledge The KG is built using entity resolution: using the graph structured, i.e. the established links, the system allows advanced queries using domain-specific entities and extracted facts. The authors argue that only the combination and integration of the sources in the KG allow such functionalities. This means only the KG makes the knowledge explicitly available via their search system.

(ii) Cotroneo et al., 2018. From Data Search to Data Showcasing: The role of Semantic Technologies in a New Service [52]

Research institutions (e.g. universities) have been struggling with finding available datasets relevant for their research. The paper discusses the creation of a research data management system, which combines cross-organisational metadata to support the search for data. The challenge in this work is that current research data repositories do not provide complete metadata, and if available, the metadata is only available as free-text. For instance, the corresponding authors’ institutions cannot directly be linked due to varying spellings, etc. Using a KG of institution, publications, and datasets the free-text metadata can be disambiguated and linked. Eventually, the graph can be used to set up a search engine over the indexed datasets.

Combination of Data Sources Elsevier’s previous datasets search engine is based on incomplete free-text labels (of institutions). In order to improve the available metadata and thus the search for datasets two kinds of semantic technologies are made use of. The paper describes the integration of metadata about publishing institutions, publications, and datasets from several external sources (e.g., DataCite) into the internal Elsevier KG “Scopus”.

Creation of new Knowledge Through the integration of the cross-organisational metadata new links between entities are created. Through the combination of explicit links between entities, existing in the individual databases, previously disconnected nodes in the newly created graph are connected. That means, in the KG e.g. institutions, publications, and datasets are
Table 2

Surveyed ISWC industry track publications.

<table>
<thead>
<tr>
<th></th>
<th>year</th>
<th>term KG</th>
<th>internal</th>
<th>external</th>
<th>both</th>
<th>use open KGs</th>
<th>knowledge creation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019</td>
<td>7</td>
<td>9</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>8</td>
<td>10</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>5</td>
<td>10</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>20</td>
<td>29</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>

connected, which generates this new knowledge, and now allows to search e.g. for publications and datasets from certain institutions. The individuals are therefore able to utilize the knowledge stored in the KG for the purpose of finding relevant datasets.

(iii) Ireson and Ciravegna, 2017. FootballWhispers: Transfer Rumour Detection [53]

Football Whispers uses social media activity (e.g. Twitter) to predict the likelihood of transfers of players between football teams. The authors rely on rumours on social media to get an indication of player transfers. However, a KG is used to identify players and teams in the noisy, informal social media posts. Available (crowdsourced) open KGs, such as DBpedia and Wikidata, are a rich source of players, teams, and even their transfer history.

Combination of Data Sources Football Whispers is based on the integration of three external sources into their KG: the knowledge base “Opta Sports”, which provides football related domain knowledge, Wikidata, and DBpedia. The entities available in “Opta Sports” are mapped to the open KGs in order to gain alternative labels (e.g. nicknames, multilingual spellings) for players, teams, etc.

Creation of new Knowledge By adding alternative labels (e.g. nicknames) for the entities the social media activity (e.g. tweets) can be filtered and analyzed based on players, teams, etc. more accurately. The collected social media posts are then analyzed and used as evidence for potential transfers.

5. Summary & Conclusions

We have herein investigated how KG technologies indeed support organisations and companies in the process of knowledge creation. How knowledge is created has been researched extensively in the KM community: We have provided a brief overview of definitions of knowledge and discussed the importance of explicit and implicit knowledge for organisations. Additionally, we have given a definition and short introduction to KGs.

On that basis, we have surveyed different implementations of KGs and analysed a set of research papers along various dimensions. In particular, we have focused on papers that describe the process of new knowledge creation through the use of KG techniques, in a company environment. To this end, we have surveyed all available publications of the industry track at the past ISWC conferences – a total of 48 papers. Overall we could identify three different categories of knowledge generation:

C1 Knowledge generation by re-structuring: We observed a first knowledge generation process already while companies re-structure information in a graph structure. The re-organisation and structuring – e.g. by using new class/type hierarchies – allows to apply graph-based techniques such as entailment, and already potentially makes new knowledge explicitly available (which was unavailable before).

C2 Knowledge generation by integration: In various applications we saw that company’s resources were combined and interlinked (using techniques such as named-entity resolution and relation extraction). These newly discovered links between the resources enable applications such as entity-based search or recommendation.

C3 Knowledge generation by enrichment: A common application of KGs is the use of external sources to enrich the company’s resources. To this end, openly available KGs such as DBpedia or Wikidata are used. These graphs provide rich, crowd-sourced, common knowledge; entities that already exist in the company’s KG can be enriched by adding additional properties (e.g. adding players’ nicknames, labels, numbers, etc., cf. Section 4.2) available in the external sources.

From a KM perspective, it can be argued that KGs could foster the externalisation process of implicit knowledge in organisations. According to Nonaka [27], externalisation plays an important role along the
knowledge generation process. Based on the idea that knowledge is defined as the capacity to act [31, 54] an improved externalisation expands the capacity to act as the level of the explicit knowledge increases.

Still, for other definitions of knowledge e.g. the justified true belief, it is questionable if a machine can truly generate knowledge in that terms. This is because some definitions require a human (action) to make sense of a situation or information provided in order to be considered knowledge. However, even following those human mind centered definitions of knowledge, there is no contradiction with our argument that KGs can facilitate the externalisation process of knowledge (i.e. making tacit knowledge explicit). Basically our understanding of knowledge only influences the point in time, when we can say that new knowledge has been created. Based on a human centered definition we acquire knowledge at the time when a human makes sense of the information externalised through the KG.

However, following the capacity to act approach, we acquire knowledge as soon as it is made explicit and provides the basis for action.

While our current literature study is limited in terms of case studies and surveyed publications, we mainly identified current reported uses of KGs to fall into three complementary categories: (i) knowledge generation by re-structuring, (ii) knowledge generation by integration, and (iii) knowledge generation by enrichment. Although we can only report here what companies published (in short, 2-pages industry track papers), we may not claim to give a full overview of current applications of KGs in organisations; still we can conclude that these applications are limited to a narrow view on how KGs can address Knowledge Creation. That is, currently surveyed use cases suggest that KGs – while contributing to some extent to knowledge creation and transfer – still have more potential, when considering this question in a more holistic manner: indeed, there are hardly any works that report and investigate yet how the applications of KGs have actually increased the capacity to act.

We claim that herein a Knowledge Management perspective, as we have introduced it in the first part of the present paper, may be worthwhile to explore. To this end, we close with a set of three directions to be taken into account in further research about Knowledge Graphs:

**D1 Actionable Knowledge**: The current focus in KG research still lies mostly in organizing rather static, factual "knowledge". In this regard we may question in how far it actually represents all facets of knowledge: what could a converging definition of knowledge between communities look like and, particularly, in how far do KGs address the capacity to act? In order to answer this question, we particularly lack definitions and examples of actionable knowledge graphs, which also encode process or behavioural knowledge, norms, etc. We note that the Semantic Web or also the Multi-Agent Systems community have made various contributions in this respect in the past; e.g. in terms of attempting to define and researching ontologies on processes, norms and policies, or on communicative action models, as well as simulations of social behaviours under considerations of social norms, etc. Yet little of this work has been transferred to the "paradigm" shift to KGs, nor yielded practical applications that could claim to have contributed to resp. knowledge creation at scale.

**D2 Embodiment of Knowledge**: The role of humans in the loop and in how far technological means (incl. KGs) can actually assist humans is not central enough, or is largely restricted to "human computation" rather than perceiving human cognition and communication as a central entity of knowledge creation itself. That is, while several works emphasise on humans in the loop for KG creation [55], we need further work on communicating and interfacing created and collected knowledge between machines and humans for making knowledge generation indeed more effective. To this end, we need to better understand how human knowledge workers communicate with and about data, and what kinds of information (often dubbed background "knowledge") and signals (again, social behaviour and processes) they leverage to achieve mutual understanding.

**D3 Organisational vs. Individual Knowledge**: Referring back to our examples, we see that the creation and transfer of knowledge needs to consider an organisational dimension: all three pro-

---

Example venues that have published articles in this area include the ongoing Conference series on Autonomous Agents & Multi-Agent Systems (AAMAS), https://dblp.org/db/conf/atal, also seemingly discontinued (though hopefully not forgotten) venues such as the workshops in Computational Logic in Multi-Agent Systems (CLIMA), https://dblp.org/db/conf/clima/, Cooperative Information Agents https://dblp.org/db/conf/cia/
totypical cases – (i) the integration of knowledge within an organisation, (ii) making knowledge explicit, accessible and actionable within and across communities and organisations, or (iii) enrichment of organisational knowledge with external/common knowledge – consider community boundaries for knowledge sharing centrally. It is therefore crucial to look at the role of KGs to enable the sharing and generation of knowledge not only at the individual level, but especially focus at the process of knowledge sharing between organizations/communities. According to Nonaka [35], the process of knowledge creation takes place at very different ontological levels (individual, group, organization, inter-organizations). Therefore, in future works, we will have to investigate more closely, at which of these ontological level KGs could be used.

When speaking about Knowledge Graphs, three dimensions should be fundamentally integrated, or – as a community – we better start looking for a more accurate term.

References


