Chapter 1
Enterprise Knowledge Graph: An Introduction

Jose Manuel Gomez-Perez, Jeff Z. Pan, Guido Vetere and Honghan Wu

```
A knowledge graph consists of a set of interconnected typed
     entities and their attributes.
```

Compared to other knowledge-oriented information systems, the distinctive features of knowledge graphs lie in their special combination of knowledge representation structures, information management processes and search algorithms. The term 'Knowledge Graph' became well known in 2012 when Google started to use knowledge graph in their search engine, allowing users to search for things, people or places, rather than just matching strings in the search queries with those in Web documents. Inspired by the success story of Google, knowledge graphs are gaining momentum in the world's leading information companies.

The idea of a knowledge graph is not completely new though. The original idea dates back to the knowledge representation technique called the Semantic Network. Later on, researchers in Knowledge Representation and Reasoning (KR) addressed

J.M. Gomez-Perez (✉)
Expert System, Prof. Waksman 10, 28036 Madrid, Spain
e-mail: jmgomez@expertsystem.com

J.Z. Pan
University of Aberdeen, King's College, Aberdeen AB24 3UE, UK
e-mail: jeff.z.pan@abdn.ac.uk

G. Vetere
IBM Italia, via Sciangai 53, 00144 Rome, Italy
e-mail: gvetere@it.ibm.com

H. Wu
King's College London, De Crespigny Park, London SE5 8AF, UK
e-mail: honghan.wu@kcl.ac.uk

J.Z. Pan et al. (eds.), *Exploiting Linked Data and Knowledge Graphs in Large Organizations*, DOI 10.1007/978-3-319-45654-6_1

some well-known issues on the Semantic Network when standardising the modern version of Semantic Network, or RDF (Resource Description Frameworks). It turns out that knowledge representation techniques, such as Knowledge Graph or Semantic Network, are useful not only for Web search, but also in many other systems and applications, including enterprise information management. The focus of the book, therefore, is about constructing, understanding and exploiting knowledge graphs in large organisations.

The basic unit of a knowledge graph is (the representation of) a singular *entity*, such as a football match you are watching, a city you will visit soon or anything you would like to describe. Each entity might have various attributes. For example, the attributes of a person include name, birthdate, nationality, etc. Furthermore, entities are connected to each other by *relations*; e.g. you *follow* one of your colleagues in Twitter. *Relations* can be used to bridge two separate knowledge graphs. For example, by saying that your Twitter ID and the ID on your driving license are denoting one and the same person, this actually interlinks Twitter data with the information space in the driver licensing agency of your country. Not surprisingly, each entity needs an identification to distinguish one another. This is the final jigsaw in the knowledge representation of knowledge graphs. Note that to facilitate the interlinking between various knowledge graphs, the entity IDs need to be *globally* unique. Types of entities and relations are defined in some machine-understandable dictionaries called ontologies. The standard ontology language is called OWL (Web Ontology Language).

The quality of a knowledge graph is crucial for its applications. For example, a knowledge graph should be consistent. In the above example, it could be the case that your contact address in your driving license is different than that in your Twitter profile. To create a knowledge graph connecting these two information spaces, such inconsistency should be resolved by keeping the correct one. In addition to consistency, one also needs to consider correctness, and coverage of knowledge graphs, as well as efficiency, fault tolerance and scalability of services based on knowledge graphs. Many of those aspects are related to, among others, the schema (ontology) of a knowledge graph.

```
A knowledge graph has an ontology as its  schema defining
    the vocabulary used in the knowledge graph.
```

1.1 A Brief History of Knowledge Graph

1.1.1 The Arrival of Semantic Networks

Knowledge management in early human history was largely shaped by oral communication before the invention of languages, which then allowed human knowledge to be recorded and passed on through generations. One of the first computer-based knowledge representation approaches are *Semantic Networks*, which represent knowledge in the form of interconnected nodes and arcs, where nodes represent objects, concepts or situations, and edges represent the relations between them, including is-a (e.g. “a chair is a type of furniture”) and part-of (e.g. “a seat is part of a chair”).

As regards the origin of Semantic Networks [38], some researchers argue that Semantic Networks have come from Charles S. Peirce’s existential graphs, while many of them pay tribute to Quillian, who was the first to introduce Semantic Networks in his semantic memory models [194]. Semantic memory refers to general knowledge (facts, concepts and relationship), such as a chair. It is different from another kind of long-term memory, i.e. episodic memory, which relates to some specific events, such as moving a chair. After Quillian, many variants of Semantic Networks were proposed.

Compared to formal knowledge representation and reasoning formalisms, such as predicate logics, Semantic Networks are relatively easy to use and maintain. On the other hand, they suffer from some limitations. For example, there is no formal syntax and semantics for Quillian’s Semantic Network. This leaves room for users to have their own interpretations of constructors in Semantic Networks, such as the is-a relation. This approach may be seen as flexible for some, but it is also criticised for making it hard to integrate Semantic Networks while preserving their original meaning. Furthermore, Semantic Networks do not allow users to define the meaning of labels on nodes and arcs.

1.1.2 From Semantic Networks to Linked Data

RDF (Resource Description Framework) is a modern standard from W3C, addressing some of the issues related to classic Semantic Networks in terms of the lack of formal syntax and semantics. For example, the is-a relation can be represented by the subClassOf property in RDF, the semantics of which is clearly defined in the RDF specifications. It should be pointed out that RDF does not address all the limitations of a Semantic Network, e.g. RDF does not allow users to define concepts either. This is, however, addressed by OWL (Web Ontology Language), a W3C standard for defining vocabularies for RDF graphs. In OWL, the part-of relation is not a built-in relation like the subClassOf property. Instead, it is a user-defined relation that can be expressed by using the existential constructor. Description Logics [18, 184] are the underpinning of the OWL standard in the Semantic Web. More details of RDF and OWL can be found in Chap. 2.

Based on RDF and OWL, Linked Data is a common framework to publish and share data across different applications and domains, where RDF provides a graph-based data model to describe objects. OWL offers a standard way of defining vocabularies for data annotations. In the Linked Data paradigm, RDF graphs can be linked together by means of mappings, including schema-level mapping (subClassOf) and object-level mapping (sameAs).

1.1.3 Knowledge Graphs: An Entity-Centric View of Linked Data

In 2012, Google popularised the term *Knowledge Graph* (KG) with a blog post titled '*Introducing the Knowledge Graph: things, not strings*',[1] while simultaneously applying the approach to their core business, fundamentally to the Web search area. Among other features, the most typical one from the user's perspective is that, in addition to a ranked list of Web pages resulting from the keyword search, Google also shows a structured knowledge card on the right, which is a small box containing a summarised information snippet about the entity that probably solves the search. Such a knowledge card contains additional information relevant to the search, contributing to relieving the burden on the user's side to pick up relevant Web pages to find answers manually. Furthermore, relations with other entities in the KG are suggested, increasing the feeling of serendipity and stimulating further exploration by the user. In most cases, such knowledge cards sufficiently fulfil searchers' information needs, significantly improving the efficiency of Web search systems both in terms of time spent per search and quality of the results.

Inspired by the successful story of Google, knowledge graphs are gaining momentum in the World Wide Web arena. In recent years, we have witnessed an increasing industrial take-up by other Internet giants, which include Facebook's Graph Search and Microsoft's Satori, continued effort made in industrial research, e.g. Knowledge Vault [69], posting community-driven events (Knowledge Graph Tutorial in WWW2015[2]; KG2014[3]), entering into academia–industry collaborations and the establishment of start-ups that specialised in areas such as Diffbot[4] and Syapse.[5] All these initiatives, taken in both academic and industrial environments, have further developed and extended the initial Knowledge Graph concept which was popularised by Google. Additional features, new insights and various applications have been introduced and, as a consequence, the notion of knowledge graphs has grown into a much broader term that encapsulates a whole line of community effort in its own right, new methods and technologies.

To explain the subtle differences between knowledge graph and Linked Data better, we first need to introduce some basic concepts. Thus, we will postpone such detailed discussions to Sect. 2.4, after providing an introduction on the background knowledge in Sects. 2.1–2.3.

---

[1] http://googleblog.blogspot.co.uk/2012/05/introducing-knowledge-graph-things-not.html.

[2] http://www.www2015.it/tutorials-19/.

[3] http://www.cipsc.org.cn/kg2/index_en.html.

[4] http://www.diffbot.com/products/.

[5] http://syapse.com/.

1.2 Knowledge Graph Technologies in a Nutshell

A knowledge graph based information system usually forms an ecosystem comprising three main components: construction, storage and consumption. Relevant knowledge graph technologies can be classified into one of these components of such an ecosystem where their contribution is most critical. As regards knowledge graph construction and storage, one finds technologies and tools for:

- knowledge representation and reasoning (languages, schema and standard vocabularies),
- knowledge storage (graph databases and repositories),
- knowledge engineering (methodologies, editors and design patterns),
- (automatic) knowledge learning including schema learning and population.

For the first three items, the majority of technologies are derived from the areas of KR, Databases, Ontologies and the Semantic Web. For knowledge learning, on the other hand, frameworks and technologies from Data Mining, Natural Language Processing and Machine Learning are typically employed.

From the consumption point of view, knowledge graphs' content can be directly accessed and analysed via query languages, search engines, specialised interfaces and/or generation of (domain/application-specific) graph summaries and visual analytics. In many other cases, a knowledge graph can enhance the effectiveness of a traditional information processing/access task (e.g. information extraction, search, recommendation, question answering, etc.) by providing a valuable background domain knowledge.

In this book, we cover knowledge graph technologies of all the above types, ranging from foundational representation languages like RDF to advanced frameworks for graph summarisation and question answering. Some of these technologies are useful for understanding knowledge graphs, while others help in exploiting knowledge graphs to support intelligent systems and applications.

1.3 Applications of Knowledge Graphs for Enterprise

Back in 2008, ongoing and future trends in semantic technologies were forecast to lie at the intersection of three main dimensions:

- natural interaction,
- the Web 2.0,
- service-oriented architectures.

If we abstract away from those particular terms, the actual meaning becomes quite simple:

- ease of *access* to computer systems by end users,

- *empowerment* of user communities to represent, manage and share knowledge in collaborative ways,
- machine *interoperability*.

Since then, countless research challenges have been faced in areas such as Knowledge Acquisition, Representation and Discovery, Knowledge Engineering Methodologies, Vocabularies, Scalable Data Management Architectures, Human–Computer Interaction, Information Retrieval and Artificial Intelligence, where semantic technologies have been involved, contributing to crucial advances in knowledge-intensive systems.

Now, like then, the value of data as the driving force behind intelligent applications remains. However, there is a new trend gaining momentum, which lies at the realisation that such a *value is directly proportional to the interlinkedness of the data* not only in complex, open-ended systems like the Web but also in specific enterprise applications based on combinations of both corporate and open data. More suited to look-up and relatedness operations, poorly formalised but highly interconnected data are becoming more popular than highly formalised but isolated datasets. The current application landscape, more oriented towards mobile and real time, is enforcing this new paradigm shift.

Google understood this very well and in 2012 started driving this trend in the industry by releasing their Knowledge Graph as a way to master such value, a large knowledge base that enhances its search engine's results with semantic-search information gathered from a wide variety of sources. Interestingly, the Knowledge Graph provides a way to connect the dots (entities) by means of explicit relations, with both entities and relations described following formal (but lightweight) models and reusing existing datasets like Freebase. After Google, other knowledge graphs arrived at the Internet scale, including those of Microsoft and Yahoo! Nowadays, it is the turn of enterprises and public administrations to leverage the Knowledge Graph concept at a corporate level in order to describe their data, enrich it by interlinking it with other knowledge bases both within and outside their environment and revitalise the development of knowledge-intensive systems on top of it.

Compared to 2008 [24], the interest in Market Intelligence and data-intensive sectors[6] and the role of knowledge graphs have increased dramatically while others, like corporate knowledge management and open government, are still there, though with slightly different foci. Next, we give an account of some selected applications that use knowledge graphs in such sectors, which will hopefully provide insight into the potential impact and future opportunities of knowledge graphs.

Corporate Knowledge Management

Open Innovation

Nowadays, especially after the recent financial downturn, companies are looking for much more efficient and creative business processes so as to place better solutions in the market in less time with less cost. There is a general impression that communication and collaboration, especially mixed with Web 2.0 approaches within companies

[6]IDG Enterprise Big Data report—http://www.idgenterprise.com/report/big-data-2.

and ecosystems (so-called Enterprise 2.0 [156]), can boost the innovation process with positive impacts on business indicators.

Open innovation [45] within an Enterprise 2.0 context is one of the most popular paradigms for improving the innovation processes of enterprises, based on the collaborative creation and development of ideas and products. The key feature of this new paradigm is that knowledge is exploited in a collaborative way flowing not only between internal sources, e.g. R&D departments, but also between external ones such as employees, customers, partners, etc. In this scenario, corporate knowledge graphs can be used to (i) support the semantic contextualisation of content-related tasks involving individuals and roles and (ii) help in discovering relations between communities of employees, customers and providers, with shared knowledge and interests.

The introduction of the open innovation paradigm in an enterprise entails not just a modification of corporate innovation processes but also a cultural change which requires support by an advanced technological infrastructure. Corporate knowledge has to be made explicit, exchanged and shared between participants, and therefore tools for knowledge management, analysis support and information structuring are required to make these tasks affordable and the knowledge available to all the involved actors. In addition, tools supporting the innovation process need to provide a high degree of interactivity, connectivity and sharing. In a scenario where collaborative work is not supported and members of a community could barely interact with each other, solutions to everyday problems and organisational issues rely on an individual's initiative. Innovation and R&D management are complex processes for which collaboration and communication are fundamental. They imply creation, recognition and articulation of opportunities, which need to be evolved into a business proposition at a subsequent stage. Interactivity, connectivity and sharing are the features to consider when designing a technological framework for supporting collaborative innovations [90]. All these characteristics can be identified in Enterprise 2.0 environments.

However, Enterprise 2.0 tools do not provide formal models which are used to create complex systems that manage large amounts of information. This drawback can be overcome by incorporating corporate knowledge graphs introducing computer-readable, interlinked representations of entities. Open innovation platforms similar to the one described in [1] leverage the concept of a corporate knowledge graph to relate people, interests and ideas in a corporate knowledge management environment throughout sectors, involving employees, clients and other stakeholders.

The impact of knowledge graphs through their application in open innovation is illustrated by their adoption in large corporations belonging to several sectors such as banking, energy and telecommunications (see further details in [45]), with companies such as Bankinter, Repsol and Telefonica, which have positioned themselves at the forefront of these efforts. What all these efforts have in common is the need to connect innovative ideas and people in order to orchestrate a healthy innovation ecosystem, addressing several challenges, like:

- Handling the information created by thousands of employees,
- evaluating their ideas efficiently,
- reducing false positives (ideas that reach the market and fail) and false negatives (valuable ideas which are rejected even before they can reach the market),
- stimulating the communication among people located around the globe, in different languages.

Intra-enterprise Micro-knowledge Management

As seen above, knowledge management is one of the key strategies that allow companies to fully tap into their collective knowledge. However, two main entry barriers usually limit the potential of this approach: (i) the barriers that employees encounter discouraging them from strong and active participation (knowledge providing) and (ii) the lack of truly evolved intelligent technologies that enable employees to easily benefit from the global knowledge provided by the companies and other users (knowledge consuming). In [188], miKrow, a lightweight framework for knowledge management, was proposed based on the combination of two layers that exploit corporate knowledge graphs to cater to both needs: a microblogging layer that simplifies how users interact with the whole system and a semantic engine that performs all the intelligent heavy lifting by combining semantic indexing and search of microblogs and users.

The miKrow interaction platform is a Web application that is designed as per the Web 2.0 principles of participation and usability. miKrow centres interaction around a simple text box user interface with a single input option for end users, where they are able to express what they are doing, or more typically in a work environment, what they are working at. This approach diverges from classical KM solutions which are powerful yet complex, following the idea of simplicity behind the microblogging paradigm in order to reduce the general entry barriers for end users. The message is semantically indexed against the underlying knowledge graph so that it can be retrieved later, as well as the particular worker linked to it. miKrow's semantic functionalities are built on top of the underlying knowledge graph, which captures and relates the relevant corporate entities.

Market Intelligence

According to the consulting company International Data Corporation (IDC) in its 2014 IDG Enterprise Big Data report, on an average enterprises spent $8M on leveraging value out of data in 2014, with penetration levels of 70 % and 56 % for large enterprises and SMEs, respectively. Improving the quality of the decision-making process (59 %), increasing the speed of decision-making processes (53 %), improving planning and forecasting (47 %), and developing new products/services and revenue streams (47 %) are the top four areas accelerating investment in data-driven business initiatives.

This trend is especially acute in the digital content and advertising sector. The communication between brands and consumers is set to explode. Product features are no longer the key to sales and the combination of both personal and collective benefits is becoming an increasingly crucial aspect. As a matter of fact, brands providing such value achieve a higher impact and consequently derive clearer economic benefits. On the other hand, millennials [98] are taking over, inducing a dramatic change in the way consumers and brands engage and what channels and technologies are required to enable the process. As a result, traditional boundaries within the media industry are being stretched and new ideas, inventions and technologies are needed to keep up with the challenges raised by the increasing demands of this data-intensive, in-time, personalised and thriving market.

HAVAS, the fourth largest media group worldwide, seeks to interconnect start-ups, innovators, technology trends, other companies and universities worldwide in one of the first applications of Web-scale knowledge graph principles to the enterprise world and media [46]. The resulting enterprise knowledge graph supports analytics and strategic decision-making for the incorporation of such talent within their first 18 months life span. Such an endeavour involves the application of semantic technologies by extracting start-up information from online sources, structuring and enriching it into an actionable, self-sustainable knowledge graph, and providing media businesses with strategic knowledge about the most trending innovations. While the previous success stories deal with the management of corporate knowledge within corporations, in this case the focus lies in creating competitive intelligence.

As we already know, innovation is often misunderstood and difficult to integrate into corporate mind-set and culture. So, why not activate relevant external talent and resources when necessary? The discovery and surveillance of trends and talent in the start-up ecosystem can be time consuming, though. HAVAS' knowledge graph sets its semantic engineering to run a surveillance monitoring of the entrepreneurial digital footprint, collecting and gathering fruitful insight and information, which provides the staff with clear leads for analysis. By automating part of the research process, analysts can get there faster and more accurately than competitors, leveraging millions of data points, and implementing consistency through a single and shared knowledge entry point. At the moment the knowledge graph is being opened to HAVAS' network, with teams in 120 offices around the world and clients, providing access to knowledge about the best-in-class talent to implement new thinking and cutting-edge solutions to the never-ending and evolving challenges within the media industry. Based on the knowledge graph, teams also rate and share experiences, ensuring that learning can be propagated across the network.

IBM Watson

IBM Watson is a cognitive computing platform available in the cloud, developed by IBM as an outcome of the *Jeopardy!* Q&A challenge[7]; cf. Sect. 7.2 and the Foreword of this book by Chris Welty. Watson uses Natural Language Processing and Machine Learning to discover insights from large amounts of unstructured data and provides a variety of services to work with this knowledge. Knowledge Graphs (such as Prismatic, DBPedia and YAGO) were at the core of the IBM's Q&A system.[8] IBM Watson services available today provide KGs capabilities through many services

[7] http://www.ibm.com/smarterplanet/us/en/ibmwatson/.

[8] IBM Journal of Research and Development, Vol. 56, No. 3/4, May/July 2012.

and application program interfaces (APIs), such as the Watson Concept Expansion and Insight.[9] Ongoing research and development aim at extending the availability of large structured knowledge bases to Dialog Services and other cognitive front ends.

1.4 How to Read This Book

1.4.1 Structure of This Book

This book introduces the key technologies for constructing, understanding and exploiting knowledge graphs. We hope you like reading this chapter so far. The rest of this book contains three parts, as illustrated in Fig. 1.1 (p. 11):

- **Part 1** contains Chaps. 2 and 3, in which we first introduce some basic background knowledge and technologies, and then present a simple architecture in order to help you to understand the main phases and tasks required during the lifecycle of knowledge graphs.
  - **Chapter** 2 introduces the background knowledge for studying and understanding the Knowledge Graph. Furthermore, we include a bit more discussion in the end to clarify the relations between Knowledge Graphs and Linked Data, as well as different purposes of building knowledge groups, e.g. for Web search versus for enterprise information systems.
  - **Chapter** 3 introduces a three-layer architecture of the Knowledge Graph application: (L1) Acquisition and Integration Layer; (L2) Knowledge Storing and Accessing Layer; and (L3) Knowledge Consumption Layer.
- **Part 2** is the main technical part for the Knowledge Graph, which contains Chaps. 4–7.
  - **Chapters** 4 **and** 5 further explain the layer L1 and address the state-of-the-art technology of knowledge acquisition and ontology construction.
  - **Chapters** 6 **and** 7 further explain the layer L3, where Chap. 6 introduces the key technologies of summarisation service, while Chap. 7 introduces the techniques of applying knowledge graphs in question answering (like the IBM Watson DeepQA).

  Based on the level of technical details, we have placed an asterisk on the titles of some chapters and sections, which contain detailed technical descriptions (e.g. formal definitions or formulas) or advanced topics (e.g. statistical/logical reasoning). Specifically, they are Chap. 5, Sects. 6.4 and 7.4.

---

[9] http://www.ibm.com/smarterplanet/us/en/ibmwatson/developercloud/.

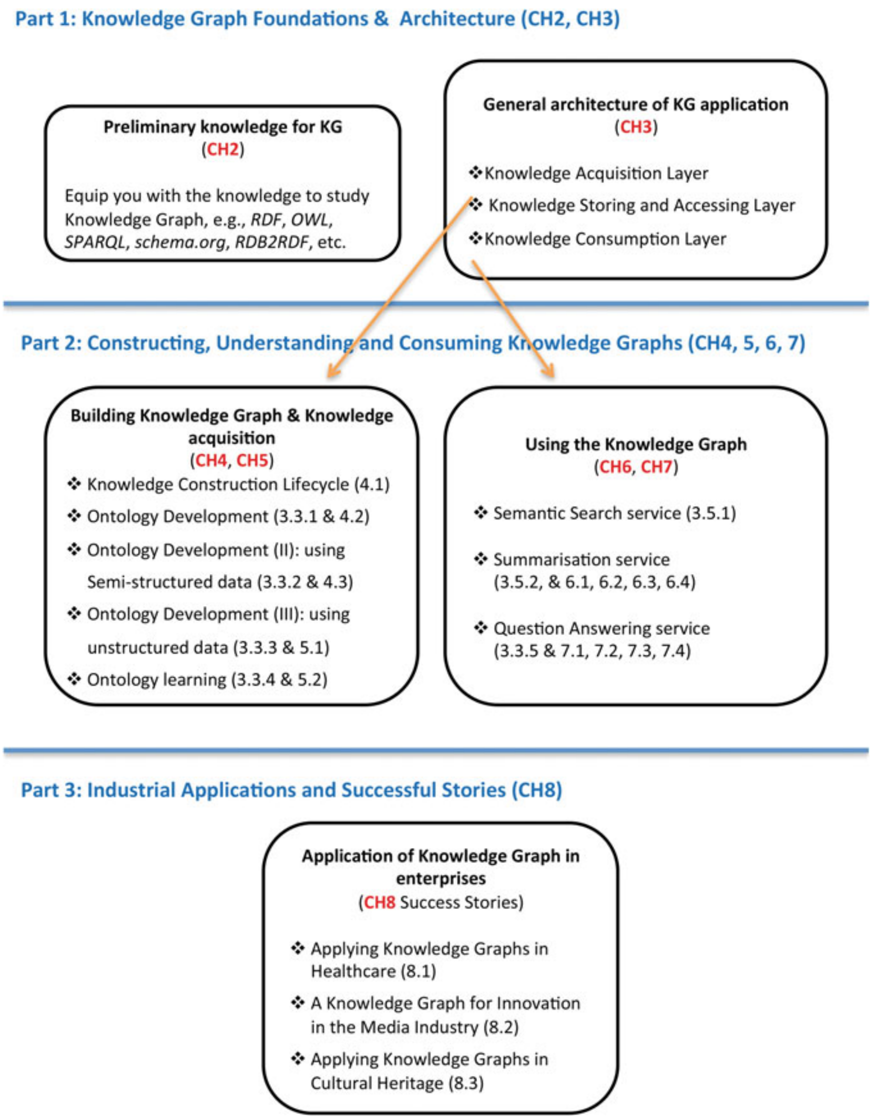

Fig. 1.1 The three parts of the main content of this book

- **Part 3** (Chap. 8) introduces the successful stories of applying Knowledge Graph in Healthcare (8.1), Media Industry (8.2) and Cultural Heritage (8.3).

In Chap. 9 we conclude this book which shares some valuable experience of the editors and authors about their works on knowledge graphs.

1.4.2 Who This Book Is For

This book is for academic researchers, knowledge engineers and IT professionals who are interested in acquiring industrial experience in using knowledge graphs for enterprises and large organisations. The book provides readers with an updated view of methods and technologies related to knowledge graphs, including illustrative corporate use cases.

I am an academic researcher/postgraduate student, what can I learn from this book?

For readers who are familiar with semantic technologies, this book provides an overview of the state of the art in knowledge graph technologies and of research methods and tools involved in building, managing and exploiting knowledge graphs. Readers will also benefit from insight and lessons learnt from the application of such approaches to different real-life problems and corporate environments.

I am an engineer or a manager in industry, what can I learn from this book?

Readers from industry will find in this book an open door to new and effective means to structure knowledge and link the different corporate assets in a way that modern organisations can exploit efficiently for a number of different purposes, including knowledge management and decision-making. Knowledge and software engineers will become familiar with the relevant techniques in the area while managers will find additional insight into how this paradigm can unlock new business opportunities to exploit both corporate and publicly available knowledge.

1.4.3 How to Use This Book

The content of this book is structured in three parts: the preliminary fundamental knowledge, the key technologies of Knowledge Graph and the applications.

Figure 1.1 (p. 11) can be used as a road-map across this book to remind the readers where they are in the journey, and help them to skip some sections, for example some sections are too technical to be of interest to general readers, and to find the most important content for them.

In the following, we provide a few details of each chapter.

Chapter 1 (Enterprise Knowledge Graph: An Introduction)

briefly explains why the editors and the authors presented this book. As mentioned in the title, this book is about how knowledge graphs are used in enterprises as knowledge management methods. In this chapter, it firstly introduces the brief evolutional history of Knowledge Graph and the key technologies used in it. Then it introduces the main applications of Knowledge Graph in enterprises. A guidance of how to read this book is also provided.

Chapter 2 (Knowledge Graph Foundations)

presents a high-level overview of the foundations of knowledge graphs. We want to introduce, in a very light way, all the concepts and basics we need for understanding and working with knowledge graphs. We start by describing how knowledge is represented and the query languages that are under the hood of the knowledge graphs. Next, we briefly present the models/vocabularies/ontologies that are needed for describing knowledge. Finally, we introduce a few basic transformation approaches from the original data source formats.

Chapter 3 (Knowledge Architecture for Organisations)

introduces a high-level overview of what is needed in order to create, maintain and exploit knowledge graphs. We realise that there is no one way of doing this for all organisations and all use cases of knowledge graphs; hence in this chapter, we introduce an abstract reference architecture that includes the main phases and tasks required during the lifecycle of knowledge graphs. For each of the phases and tasks, we then present a more detailed description of possible approaches, methodologies and tools which have been reported in the literature. By the end of this chapter you should have a good idea of the tasks you would likely encounter when building and maintaining knowledge graphs. You would also have a better understanding of how knowledge graphs can be used within large organisations.

Chapter 4 (Construction of Enterprise Knowledge Graphs (I))

as well as Chap. 5, focuses on the *Acquisition and Integration Layer* of Chap. 3's reference architecture. In particular, we start with a generic lifecycle for constructing and maintaining knowledge graphs in Sect. 4.1 and, then, we elaborate on the knowledge graph construction approaches a.k.a. the modelling and data lifting steps in the lifecycle. In this chapter, we focus on supervised approaches to constructing new knowledge, i.e. approaches involving human effort. Specifically, for the modelling step we introduce a competency question based ontology authoring framework, while for data lifting we discuss a semi-automated approach for creating linkages among heterogeneous data sources.

Chapter 5 (Construction of Enterprise Knowledge Graphs (II))

continues with the *Acquisition and Integration Layer* of Chap. 3 on a reference architecture, focusing on knowledge graph construction techniques. Nevertheless, we shift from semi-automated approaches to automated approaches of knowledge graph construction by describing two additional frameworks, one for entity/scope resolution of textual data (Sect. 5.1) and one for the learning of ontological schemas from data (Sect. 5.2).

Chapter 6 (Understanding Knowledge Graphs)

identifies and introduces a set of techniques that make knowledge graphs directly available to end users. Among others, we lay a special focus on knowledge graph understanding techniques, many of which were especially designed for scenarios in large organisations.

Chapter 7 (Question Answering and Knowledge Graphs)

This is a "star" chapter. We primarily consider the tasks of question answering over text documents (Sect. 7.1) and knowledge graphs (Sect. 7.2), and we present an overview of relevant methodologies, technologies and systems. Moreover, in Sect. 7.3, we describe a state-of-the-art question-answering system that combines knowledge coming from the text analysis and knowledge graphs.

Chapter 8 (Success Stories)

presents success stories of the applications of Knowledge Graph techniques from various domains (healthcare, media and culture) and different organisations (international company—IBM, Small and Medium Enterprises—HAVAS, and University—the University of Aberdeen).

Chapter 9 (Conclusion and Outlook)

concludes the book with a brief review of the whole book. Furthermore, it shares the valuable experience from the editors and authors of "things to keep in mind" when adopting knowledge graphs.

We hope you will enjoy "Exploiting Linked Data and Knowledge Graphs in Large Organizations", as we have been enjoying it.

http://www.springer.com/978-3-319-45652-2

Exploiting Linked Data and Knowledge Graphs in Large Organisations
Pan, J.Z.; Vetere, G.; Gomez-Perez, J.M.; Wu, H. (Eds.)
2017, XVIII, 266 p. 59 illus., 44 illus. in color., Hardcover
ISBN: 978-3-319-45652-2