Chapter 1
Introduction: What Is a Knowledge Graph?

Knowledge graphs are critical to many enterprises today: They provide the structured data and factual knowledge that drive many products and make them more intelligent and magical. (Noy et al. 2019)

Abstract Since its inception by Google, Knowledge Graph has become a term that is recently ubiquitously used yet does not have a well-established definition. This section attempts to derive a definition for Knowledge Graphs by compiling existing definitions made in the literature and considering the distinctive characteristics of previous efforts for tackling the data integration challenge we are facing today. Our attempt to make a conceptual definition is complemented with an empirical survey of existing Knowledge Graphs. This section lays the foundation for the remainder of the book, as it provides a common understanding on certain concepts and motivation to build Knowledge Graphs in the first place.

1.1 Introduction

Smart speakers such as Alexa and Google Home introduced AI-based (AI) communication means in millions soon billions of households. Even the web is currently reinventing itself by applying schema.org[1] (Guha et al. 2016). Data, content, and services become semantically annotated, allowing a software agent, so-called bots,[2] to search through the web understanding its content. Therefore, it becomes increasingly important for information, service, and product providers to be highly hearable and visible in these new online channels to ensure their future economic maturity.

[1] https://www.schema.org/

[2] https://en.wikipedia.org/wiki/Internet_bot

D. Fensel et al., *Knowledge Graphs*, https://doi.org/10.1007/978-3-030-37439-6_1

The development of such automatic methods for speech recognition is an important prerequisite for the development of automated dialog systems.[3] Their breakthrough in automatic language understanding is based on Big Data[4] and machine learning (Goodfellow et al. 2016). However, for answering a query or for running a goal-oriented dialog, more is needed. For giving a meaningful answer, an agent needs knowledge. Therefore, Google started in 2012 to develop a so-called Knowledge Graph,[5] which should contain significant aspects of human knowledge found semantically annotated on the web or in other data sources. Meanwhile a kind of hype has arisen around this technology.[6] In consequence, it becomes necessary to better understand what Knowledge Graphs are about. We approach this question complementarily. First, we try to give a conceptual answer by analyzing the underlying principles of a Knowledge Graph. Second, we provide an empirical survey on existing Knowledge Graphs.

1.2 A Conceptual Definition of Knowledge Graphs

Size matters (unknown author, most likely male)

Ehrlinger and Wöß (2016) provides a very useful and concise survey on potential definitions of Knowledge Graphs illustrating their variations. They also add a new definition which is centered on using Ontologies and reasoners deriving new knowledge. From our point of view, this definition is too exclusive and too much focused on specific methods. Let us start with a few thoughts on potential *definitions* of a Knowledge Graph. Hermeneutically, we could first distinguish the two terms constituting this concept since we have two quite different beasts.

A "*graph* is a structure amounting to a set of objects in which some pairs of the objects are in some sense related".[7] Strictly spoken we need to slightly extend this definition to multi-sets since the same object can syntactically and semantically appear several times in our graph. Some normalization can get rid of this issue, but this already implies certain specific processing techniques. This simple definition can be extended in various directions and we end up with an entire zoo of graph types: simple graphs, undirected versus directed graphs, oriented graphs, mixed graphs, multigraphs, Quiver, weighted graphs, half-edges and loose-edges graphs,

[3] https://en.wikipedia.org/wiki/Dialogue_system

[4] https://en.wikipedia.org/wiki/Big_data

[5] A. Singhal: Introducing the Knowledge Graph, things, not strings. Blog post at http://googleblog.blogspot.co.uk/2012/05/introducing-knowledge-graph-things-not.html, 2012.

[6] Just to mention a few books (Chen et al. 2016; Croitoru et al. 2018; d'Amato and Theobald 2018; Ehrig et al. 2015; Li et al. 2017; Pan et al. 2017a, b; Qi et al., 2020; Qi et al. 2013; Van Erp et al. 2017). See also Bonatti et al. (2019).

[7] https://en.wikipedia.org/wiki/Graph_(discrete_mathematics).