

Knowledge graphs: Construction, management and querying

Editorial

Mayank Kejriwal^{a,*}, Juan Sequeda^b and Vanessa Lopez^c

^a *Information Sciences Institute, University of Southern California, Los Angeles, CA, USA*

E-mail: kejriwal@isi.edu

^b *data.world, Austin, TX, USA*

E-mail: juan@data.world

^c *IBM, Dublin, Ireland*

E-mail: vanlopez@ie.ibm.com

Keywords: Knowledge graphs, knowledge graph querying, knowledge graph construction, applications, linked data, knowledge graph management

Graphs have been one of the fundamental data models in Artificial Intelligence and Databases, but the astounding growth of the Web has brought them to the forefront. Graph models and databases have always served as an interesting frontier of research in the database and data mining communities; however, with the advent of knowledge graphs, we are now starting to see Web-scale commercial applications that are directly pertinent to research that has long been a staple in the Semantic Web and the widespread use of knowledge graphs across diverse data-driven industries, changing the way information is being produced and consumed both by users and software programs.

A knowledge graph, as most practitioners now understand the term, is understood best as a model about ‘things not strings’ i.e. the ‘elements’ in knowledge graphs, whether entities, relationships or even literal attributes, have very real semantics associated with them. Some of the semantics are ontologized and formal, but even when this is not the case, exciting new research in the area of ‘representation learning’ or embeddings has revealed that KGs, like ordinary words in natural language, have implicit semantics largely shaped by their context. These embeddings have unleashed a wave of applications, some of which are unsupervised or require few explicit labels. Further-

more, as several papers accepted to this special issue illustrate, the kinds of ‘elements’ included in modern state-of-the-art KGs now include not just entities and relations, but complex ‘higher-order entities’ such as events. Many of the KGs are domain-specific, and at least one involves a non-English language.

Long before the Google Knowledge Graph, or even the popularity of the now ubiquitous phrase ‘knowledge graph’, in the Semantic Web, we have a history of being consistently focused on combining knowledge and data at scale by using graphs and semantics as a means to integrate, unify, interpret, augment and query data from diverse and multiple sources at web scale. Linked Open Data (LOD) is the most well-known ecosystem serving as a testament to this focus. Starting from just a handful of datasets in 2007, the LOD initiative has grown exponentially, on average, in the last ten years and spawned scores of applications and start-ups. Several major organizations, such as the New York Times, have leveraged Linked Open Data principles to publish their own concepts. Arguably, the LOD initiative was one of the earliest instances, using Web protocols and publishing standards, to publish what we now call knowledge graphs as ‘linked data’.

Perhaps the most important trend that the collection of papers accepted in this special issue show (and that is also reflected in the papers that were submitted to the issue) is the increasing confluence of different areas of

* Corresponding author. E-mail: kejriwal@isi.edu.

AI research under the umbrella of constructing, cleaning, querying, and building applications over, knowledge graphs. Admittedly, this is a trend that we anticipated based on our collective experiences in the recent editions of Semantic Web and World Wide Web conferences such as ISWC, ESWC and the WebConf. In fact, our own experiences are quite diverse, spanning traditional academia, industrial research labs and startups.

We have not been disappointed in our expectations for this issue. This issue has confirmed to us that no longer are knowledge graphs exclusive as a research area to the Semantic Web; indeed, the collection of accepted papers shows strong synergies and

contributions from work in natural language processing, knowledge discovery, representation learning and databases. We believe that this is a sign of research yet to come, as with more technological diffusion and interdisciplinary forums such as this special issue, more areas will start to come together in the pursuit of solving grander challenges. To that end, we would like to thank all contributors for their outstanding submissions, and our guest editorial board for painstakingly providing reviews and productive comments at every stage of the yearlong process of putting together this journal. We hope that you enjoy reading this special issue as much as we enjoyed putting it together.