From Knowledge Graphs to Knowledge Practices: On the Need for Transparency and Explainability in Enterprise Knowledge Graph Applications

Christine T. Wolf
IBM Research - Almaden
San Jose, CA, USA
ctwolf[at]us[dot]ibm[dot]com

ABSTRACT

Transparency and explainability are important concerns in building and implementing cognitive technologies, especially those intended to augment and transform domain knowledge practices. Knowledge graphs are powerful component technologies leveraged in a number of business applications, yet relevant information about their construction, refinement, and maintenance are often not visible in resulting end-use applications. Information about a graph's training set(s), for example, can be useful to understand how it may represent incomplete coverage of a domain area; information about the strength or confidence of a given association, as another example, can also help signal the need for further action or investigation. In this position paper, we explore these issues and discuss an enterprise domain use case involving sales operations. Increasing transparency and explainability in knowledge graphs can help business users better assess potential biases and appropriately integrate graph outputs into their workplace knowledge practices.

CCS CONCEPTS

• Human-centered computing

KEYWORDS

Knowledge graphs; transparency; explainability; knowledge practices; engaged interaction; enterprise

ACM Reference format:

Christine T. Wolf. 2020. From Knowledge Graphs to Knowledge Practices: On the Need for Transparency and Explainability in Enterprise Knowledge Graph Applications. Position paper for Knowledge Graph Bias Workshop at *Automated Knowledge Base Construction* (AKBC'20). *3 pages.*

1 Introduction

Graphical representations of knowledge, around since the early days of digital computing, are important building blocks for a community's communication, collaboration, and knowledge practices [8]. Knowledge graphs (KGs) are often component technologies in domain applications, for example serving as the underlying knowledge base for a Q&A system in the clinical health domain. Knowledge graphs also hold great promise for enterprise contexts, where business processes often span across a number of large, distributed, and fragmented data systems [1]. Although powerful, KGs have trade-offs: "whichever approach is taken for constructing a knowledge graph, the result will never be perfect" [7] (p. 1). There are a number of refinement techniques and methods to sharpen and expand KGs, for example completion [7]. Another concern is KG maintenance, as changes in the underlying domain knowledge structure evolve [5].

As the use of cognitive technologies becomes more widespread, AI ethics has become a pressing concern [4]. Transparency is a central issue in AI ethics, defined as "efforts to increase explainability, interpretability or other acts of communication and disclosure" [4] (p. 391). Motivations for transparency in AI include improving AI, minimizing harm, complying with regulations, fostering trust, supporting dialogue and participation, and furthering principles of democracy [4] (p. 391). Transparency also plays a role in shaping interactions and supporting human-AI *collaboration* (humans and AI working together) that goes beyond *coordination* (humans and AI working in parallel) [12]. KG explainability is a growing research topic, with a number of techniques introduced in the past few years aiming to provide users with reasoning on sequential paths and associations to infer their logical/casual rationale [6, 11, 15]. In this position paper, we raise attention to different elements of KG explainability, provoking fodder for discussion at the workshop. Drawing on an enterprise use case in the sales operations domain, we highlight two explainability areas that shape the ability of business users to incorporate KG insights into their everyday knowledge practices: *knowledge sources* (what is the evidence for this association?) and *knowledge currency* (is this association up to date?). We elaborate on these issues below.

2 Enterprise Case: Quote-to-Cash (Q2C) Practices

Recent decades have seen an explosion in digital technologies like Big Data, cloud computing, and AI, which offer great potential for enterprise transformation, often called the "cognitive enterprise" [2]. Knowledge graphs, or ontologies, are desirable for enterprise settings for a number of reasons. One is that they offer the ability to capture and make persistent organizational knowledge across process-functions and departments [1]. Enterprise knowledge graphs can also help in increasing efficiency by reducing the amount of time that workers must manually look for information

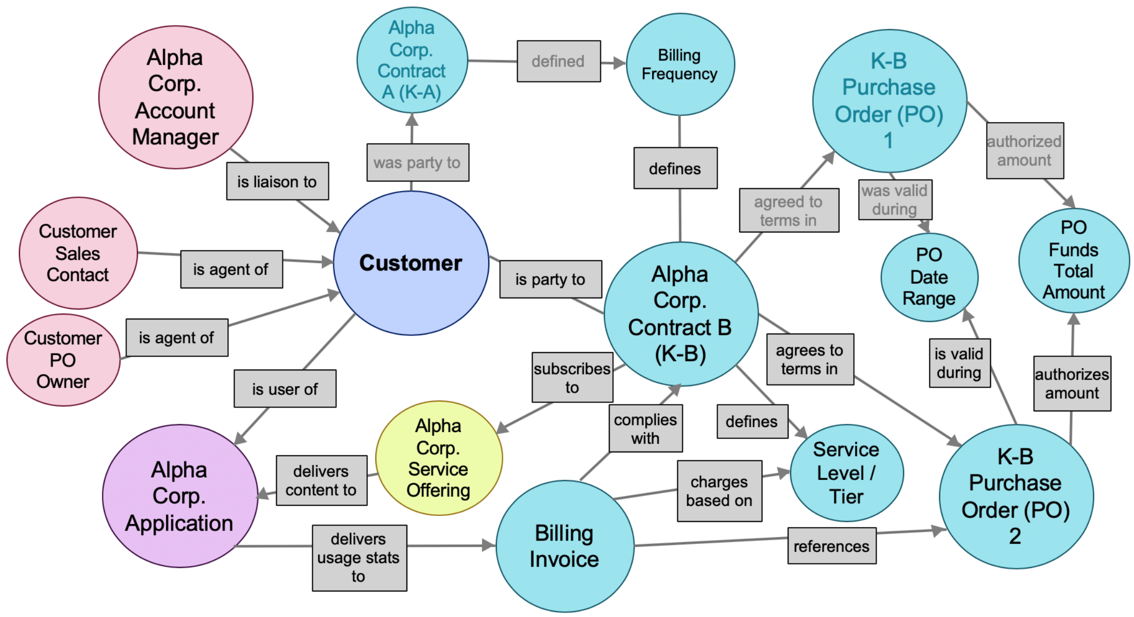

Figure 1: A knowledge graph diagram showing sample entities and associations in the Q2C invoicing business process.

across IT systems. The use of knowledge graphs, though, also raises a number of questions of transparency and explainability in human-computer interaction. For example, an enterprise Q&A system powered by a KG can provide an integrated and global view of the organization's knowledge on a given topic or domain. But depending on the specific organizational practice, business users will need to understand the source (what is the training dataset or evidence for this triple?) and currency (for changing relations, is this triple up to date?) of the information presented to them. This contextual information is necessary to ensure actionability – in order for a business user to confidently take action on the graph's outputs, KG transparency is needed. To illustrate these points, we draw on ongoing research in the quote-to-cash (Q2C) space. Q2C is an industry term of art for the overarching business process in sales organizations that encompasses all the sub-processes involved from generating a sales quote for a customer to realizing revenue or "cash" from that customer [9] This complex and integral back-office process is ripe for cognitive transformation, efforts that have been called "intelligent sales" [13]. We focus on one sub-process of the overall Q2C process: invoicing a customer. Below, we discuss how *knowledge source* and *knowledge currency* are important explainability requirements [14] in this use case.

Invoicing involves many actors – people (e.g., invoice processors, account managers), IT systems (e.g., CRM, invoicing system, purchase order system, accounts receivable), and data in various documents (e.g., contract, purchase order, invoice). Invoice processers must verify several pieces of information by looking at documents in a number of IT systems, to ensure an invoice references necessary information, such as the correct, active purchase order (PO) number. The PO is an instrument used in many business-to-business (B2B) transactions that sets aside funds to pay for contracted goods and services.

For example, Alpha Corp.[1] is an IT services vendor in our sample KG diagram (see Figure 1). Consider a typical scenario we have encountered in our ongoing work: Alpha enters into a three-year contract (K-B) with a Customer for Alpha Services. The Customer has a history with Alpha (reflected in a previous contract, K-A). Although K-B has a three-year term, the Customer will only issue a yearly PO because this cycle aligns with their fiscal budgetary interval. Thus, a new PO must be created each year in K-B's three-year term.

When preparing an invoice, processers must verify: 1) the Customer is PO-driven; 2) the PO is active and valid (not expired) and the correct PO number is reflected in the invoice; 3) any relevant usage data from the IT application (e.g., subscription overages) is reflected in the invoice; and 4) there are sufficient funds remaining in the PO to cover the invoice amount. Completing these steps involves interacting with different systems: 1) Customer details reside in a directory; 2) PO details reside in a PO-specific database, where PDF copies of the PO are located; 3) overage information comes from the delivery team who run and monitor the applications used by the Customer; and 4) PO's remaining amount involves coordinating with accounts receivable (AR) department to confirm the Customer's previous payments from the PO. For quality assurance, invoice processors must verify information by referencing forms of evidence such as a PDF copy of the PO signed by the Customer.

Q2C processes are promising candidates for KGs, used in conjunction with AI technologies, to structure and organize the knowledge about the Customer in Alpha's various systems. For example, given that the majority of the content in this graph (Figure 1) comes from various unstructured enterprise documents, (e.g., contracts and PO PDFs)., NLP (e.g., text extraction and classification) would eliminate the need for manual efforts. The KG could then be used in a Q&A dialogue system, where business

[1] This is a pseudonym.

users can run a number of domain-driven queries, enhancing a number of knowledge practices across Q2C processes. Furthermore, new associations between enterprise entities could be learned, enabling the "cognitive enterprise." Yet, requirements analysis in this use case reveals two key issues of transparency needed for our business users to be able to act upon KG outputs. The first is visibility into the *sources of knowledge* reflected in the KG. *Does this incorporate all or only some of the various IT systems they use in the invoicing process? What evidence supports this association?* The second is awareness into the currency or up-to-dateness of the KG's universe. *Does this reflect the latest PO information? How do I know when it was last updated?*

3 Summary: From Knowledge Graphs to Knowledge Practices

We conclude by posing two interlinked, conceptual questions: what are KGs for? what is transparency for? In this enterprise use case, both are closely tied to the everyday knowledge practices of enterprise business users. In order to meaningfully enhance these knowledge practices by supporting engaged and informed interaction of our business users, transparency is a key requirement of our cognitive system through visibility into its knowledge sources and the currency of that knowledge. As KGs are used in a wider array of high-stakes domain settings (e.g., financial fraud [16] or hiring decisions [3]), transparency and explainability requirements will become more and more pressing concerns, which we hope to discuss further at the workshop.

ACKNOWLEDGMENTS

Thank you to informants for sharing their time and insights into their invoice processing work practices. Thanks also to reviewers. All opinions expressed herein are my own and do not reflect institutional endorsement.

REFERENCES

[1] Gomez-Perez, J.M. et al. 2017. Enterprise Knowledge Graph: An Introduction. *Exploiting Linked Data and Knowledge Graphs in Large Organisations*. J.Z. Pan et al., eds. Springer International Publishing. 1–14.

[2] IBM Institute for Business Value 2018. *The Cognitive Enterprise: Reinventing your company with AI*.

[3] Jia, S. et al. 2018. Representation of Job-Skill in Artificial Intelligence with Knowledge Graph Analysis. *2018 IEEE Symposium on Product Compliance Engineering - Asia (ISPCE-CN)* (Dec. 2018), 1–6.

[4] Jobin, A. et al. 2019. The global landscape of AI ethics guidelines. *Nature Machine Intelligence*. 1, 9 (Sep. 2019), 389–399. DOI:https://doi.org/10.1038/s42256-019-0088-2.

[5] Khattak, A.M. et al. 2013. Change management in evolving web ontologies. *Knowledge-Based Systems*. 37, (Jan. 2013), 1–18. DOI:https://doi.org/10.1016/j.knosys.2012.05.005.

[6] Ma, W. et al. 2019. Jointly Learning Explainable Rules for Recommendation with Knowledge Graph. *The World Wide Web Conference* (San Francisco, CA, USA, May 2019), 1210–1221.

[7] Paulheim, H. 2016. Knowledge graph refinement: A survey of approaches and evaluation methods. *Semantic Web*. 8, 3 (Dec. 2016), 489–508. DOI:https://doi.org/10.3233/SW-160218.

[8] Ribes, D. and Bowker, G.C. 2009. Between meaning and machine: Learning to represent the knowledge of communities. *Information and Organization*. 19, 4 (Oct. 2009), 199–217. DOI:https://doi.org/10.1016/j.infoandorg.2009.04.001.

[9] Sinha, S. 2018. Intelligent sales lead to cash. *The Future of Customer Engagement and Experience*.

[10] Tanggaard, L. 2013. The sociomateriality of creativity in everyday life. *Culture & Psychology*. 19, 1 (Mar. 2013), 20–32. DOI:https://doi.org/10.1177/1354067X12464987.

[11] Wang, X. et al. 2019. Explainable Reasoning over Knowledge Graphs for Recommendation. *Proceedings of the AAAI Conference on Artificial Intelligence*. 33, 01 (Jul. 2019), 5329–5336. DOI:https://doi.org/10.1609/aaai.v33i01.33015329.

[12] Wolf, C. and Blomberg, J. 2019. Evaluating the Promise of Human-Algorithm Collaborations in Everyday Work Practices. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW (Nov. 2019), 143:1–143:23. DOI:https://doi.org/10.1145/3359245.

[13] Wolf, C.T. 2020. AI Ethics and Customer Care: Some Considerations from the Case of "Intelligent Sales." *Proceedings of the 18th European Conference on Computer-Supported Cooperative Work* (2020).

[14] Wolf, C.T. 2019. Explainability scenarios: towards scenario-based XAI design. *Proceedings of the 24th International Conference on Intelligent User Interfaces, IUI 2019, Marina del Ray, CA, USA, March 17-20, 2019* (2019), 252–257.

[15] Xian, Y. et al. 2019. Reinforcement Knowledge Graph Reasoning for Explainable Recommendation. *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval* (Paris, France, Jul. 2019), 285–294.

[16] Zhan, Q. and Yin, H. 2018. A loan application fraud detection method based on knowledge graph and neural network. *Proceedings of the 2nd International Conference on Innovation in Artificial Intelligence* (Shanghai, China, Mar. 2018), 111–115.