How to Effectively Make and Use Knowledge Graphs Through Collaborative Activities: A Socio-Technical Perspective

Yanan Gao, Xiyan Lv, Wenchi Ying

How to Effectively Make and Use Knowledge Graphs through Collaborative Activities: A Socio-Technical Perspective

Yanan Gao, Xiyan Lv, Wenchi Ying(✉)

School of Economics and Management,
Beijng Jiaotong University, Beijing, China
{19120605, lvxiyan, wcying}@bjtu.edu.cn

Abstract. Knowledge graphs (KG) as emerging technology contribute to knowledge networking and efficient knowledge collaboration. Sound knowledge about why organizations should develop KGs and how they can make and use KG is the key to success. However, the organized work on the Make and Use of KG can be perceived as another kind of highly collaborative activities between organizations, people, machines and systems, which received few attentions in research. Thus, we adopt socio-technical systems (STS) perspective to review, practice and study how to effectively make and use KGs through collaborative activities. This study reveals a collaborative-activity framework of KG's Make and Use and corresponding collaborative mechanisms, which may contribute to the existing literature of KG/AI, knowledge collaborative networks and STS. Organizations can use this framework to develop their own domain-specific KG and KG-based information.

Keywords: Knowledge graphs(KG) · Make and Use · Collaborative activities · Collaborative networks · Socio-technical systems

1 Introduction

Knowledge graphs (KG), first proposed by Google, is an emerging digital technology that can enhance the results of search engines and the effectiveness of cognitive artifact intelligence with information gathered from a variety of sources [1]. The concept of KG was first proposed by Google, where KG is a semantic graph consisting of nodes and edges. The nodes represent entities or concepts, and the edges represent various semantic relationships between entities or concepts [2]. In particular, based on co-occurrence analysis, social network analysis and other basic theory, KG is built to explain the structures of domains [3]. KG was originally intended to enhance the user search experience and later widely used in intelligent question answering, personalized recommendation and other fields, and further in-depth combined with the specific business needs to provide comprehensive knowledge services.

Gartner, the leading research and advisory company, highlights knowledge graphs as an emerging technology with significant impacts on business, society and people [1], which has become a hot topic in academia. Digital technologies, especially in China, has been deeply applied to and influenced many industries. Many enterprises and universities are intended to adopt KG to enhance or transform their information systems (IS) [4]. As the keystone technology of cognitive artificial intelligence, KG enables IS to contribute to knowledge sharing and intelligent collaboration among people from all walks of life, and even enables interactions between humans and machines. Thus, effective KG applications consist of the high-quality KG and the KG-based IS [4]. Obviously, the success of KG applications relies on actions in both the *Make and Use of KG* and collaboration among the actions [5].

Ongoing work on knowledge graphs can be organized in such a way that they are perceived as highly collaborative activities, and thus KGs will be broadly accepted as central knowledge hubs to overcome resistance within an organization against external knowledge [1]. Therefore, the process of KG's Make and Use is a series of interdisciplinary and cross-domain collaborative activities among social and technical entities such as organizations, people, machines, and systems [6]. However, most extant research focused on the general KG (e.g., Wikipedia and Baidu Baike) [7], computer technology level (e.g., modeling and algorithms for KG's fusion and reasoning) [2] or some domain-specific KG applications and results (e.g., specific domains of Medicine, Law and Education) [8]. Because of complexity in, and differentiation among specific domains, the process of KG's Make and Use received few attentions, which leads to the lack of experiences and frameworks to guide reusable actions for development of KG applications.

Thus, we adopt the socio-technical systems (STS) perspective as the guiding theoretical lens to address the gap. STS theory is a well-established perspective in IS research and leveraged to study the phenomena (e.g., KG applications) combining social and technical factors [9], especially technical-induced organizational tasks and actions [10]. Thus, we derive our research question: *How to effectively make and use knowledge graphs through collaborative actions from socio-technical perspective?*

According to reviews on KG literature and our practice on development of KG-based applications in an IT/IS consulting domain. We have completed preliminary findings and propose *a collaborative-activity framework of KG's Make and Use* which combines the socio-technical factors including structures, actors, technologies and tasks. The study identifies two tasks of KG applications – Make and Use, and further reveals collaboration mechanisms regarding socio-technical actions not only within the Make and the Use separately but also between them through three interaction cycles. Thus, our findings may make both important theoretical and practical contributions. First, this study contributes not only to the KG/AI and knowledge collaborative networks literature but also to the STS literature. Second, organizations can use our findings to develop their own domain-specific KG and KG-based IS, and by doing so, increase the success rate of their KG applications.

2 Research Background

2.1 Knowledge Graph (KG) and Collaboration in KG's Make and Use

Data are symbols that represent the properties of objects and events, information consists of processed data, and knowledge is the appropriate collection of information that can answers the *what/how/why* question [11]. Knowledge graph (KG) can link the different data/information streams in an intelligent and dynamic way [1]. KG has the characteristics of huge scale, rich semantics, excellent quality and friendly structure that are different from the traditional semantic network [12]. KG can be used to explicitly capture requirement semantics that are limited towards traditional databases [13]. KG can link the different streams and multimodal data in a structured way. Through in-depth semantic analysis and mining, with the help of powerful semantic processing capabilities and open interconnection capabilities, KG provides users with intelligent search and other services through a visual interface [8]. Thus, KG is an emerging digital technology that can enhance the results of search engines and the effectiveness of domain-specific cognitive AI with information gathered from a variety of sources [12].

As the keystone technology of cognitive artificial intelligence, KG enables IS to contribute to knowledge sharing and intelligent collaboration among people from all walks of life, and even enables interactions between humans and machines. Extant research mainly focused on the general KG [7], computer technology level [2] or some domain-specific KG applications and results [8]. However, research on how to build KG-based applications effectively remain limited. Gartner Hype Cycle shows that KG has still appeared on the stage of "innovation trigger" [14], which means that KG will be affected by uncertain factors on collaborative-network robustness, resource constraints, and actor selection [15] as well as the differentiation of domain-specific scenario and business needs [16]. All above lead to the lack of experiences and frameworks to guide reusable actions for development of effective KG applications. Simultaneously, effective KG applications consist of the high-quality KG and the KG-based IS [4]. Obviously, the success of KG applications relies on actions in both the *Make and Use of KG* and collaboration among the actions [5]. Organizations are eager to explore that how KG is designed and built (the Make of KG), how KG is used and embedded into concrete IS (the Use of KG) and how the Make and Use can be effectively collaborative, and the relevant research gap remains.

Furthermore, Ongoing work on knowledge graphs can be organized in such a way that they are perceived as highly collaborative activities [1], and the complexity of domain-specific KG reflects in collaborative network factors including organizations, people, machines, and systems [6]. Obviously, the process of KG's Make and Use is a series of interdisciplinary and cross-domain collaborative activities whose factors are almost consistent with the social and technical entities of socio-technical perspective [17]. Thus, socio-technical systems (STS) theory provides us with a suitable theoretical lens to address the research gap.

2.2 A Socio-Technical Perspective toward Collaboration in KG Make and Use

STS theory is a well-established perspective in IS research and leveraged to study the technical-induced organizational tasks and actions [17]. Collaborative activities in KG's Make and Use are a typical IS phenomenon so that the Make and the Use are not only technical tasks but also social tasks [17]. E.g., Andreas Blumauer [1] suggests that KG cannot be Made without support throughout the whole organizations, while new external roles/persons with diverse skills and knowledge should be introduced as well in order to support this transformation. In addition, the Use of KG must also start with planned goals and strategies which requires a series of criteria and mechanisms involving actions on workflow, skills, technological tools or platforms.

STS theory distinguishes the social system and the technical system through four entities [17] that can exactly correspond to the four factors of collaborative networks. The former involves socio-technical entities *structures* and *actors*, while the latter involves socio-technical entities *technologies* and *tasks* (See Fig.1). Specifically, *structures* refer to the KG's project organizations, institutional arrangements and working criteria; *actors* are the participants of KG projects who have different expertise or capabilities and play different roles in activities; *technologies* refer to the technological tools or platforms involved in the Make or Use of KG; *tasks* are the processes required to achieve goals or provide deliverables.

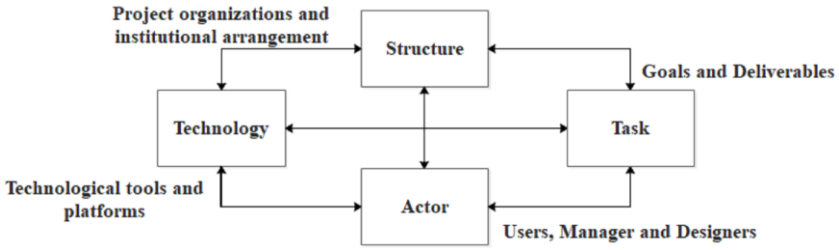

Fig. 1. The socio-technical system model (according to [9])

Obviously, the processes of Make and Use are two typical types of KG tasks that should be supported by actions from other three socio-technical entities. Thus, STS is not static but dynamic. To fulfill the tasks of KG's Make and Use, it is necessary to further identify and leverage structure-level, actor-level and technology-level collaborative actions [10]. More importantly, the interactions between Make and Use concern both rigor and relevance of KG study as well as effectiveness of KG application [5]. Simultaneously, the Make and Use of KG are not achieved at one stroke, while the interactions between the Make and Use of emerging technologies are a long-term process of alignment on adaptation, experimentation and actualization [18]. However, the socio-technical perspective on the Make and Use of emerging digital technologies also received few attentions. Thus, we derive our research question: *How to effectively make and use knowledge graphs through collaborative actions from socio-technical perspective?*

3 Research in Progress and Preliminary Findings

We have systematically reviewed literature on KG's Make and Use. Our KG project that started from 2019 is in progress and supported by HuaQiCT Co. Ltd., a class-A qualification IT/IS consulting company in China.

3.1 Collaborative-Activity Framework of KG's Make and Use

According to reviews on literature and our practice on development of KG-based applications in an IT/IS consulting domain, we have completed preliminary findings and propose *a collaborative-activity framework of KG's Make and Use* (See Fig. 2) which combines the socio-technical factors including structures, actors, technologies and tasks.

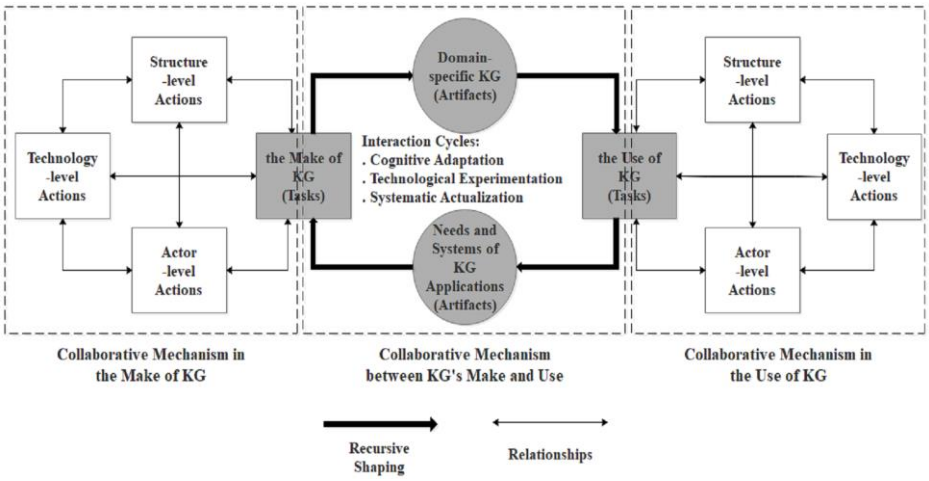

Fig. 2. Collaborative-activity Framework of KG's make and use

The study identifies two tasks of KG applications – Make and Use, and further reveals collaboration mechanisms regarding socio-technical actions not only within the Make and the Use separately but also between them. Thus, the focal components in collaborative-activity framework of KG's Make and Use are showed in Table 1.

Table 1. Focal components in collaborative-activity framework of KG's Make and Use

Tasks	*The Make of KG* • The expertise-dominant task • The technical-dominant task	*The Use of KG* • The graph-visualization task • The graph-association task
Structures	• Technical-oriented group • professional crowdsourcing community	• Product/application-oriented group • IT/IS-oriented group
Actors	• KG architects with the ambidextrous skills on knowledge management and big data analytics • Technical experts mastering modeling, algorithm and other information technologies of KG • Domain-specific experts familiar with the business context • Domain-specific employees for annotation and labeling	• Product managers (PM) / Business analysts (BA) • Domain-specific experts familiar with the business context • Technical experts mastering KG-based IT/IS skills • KG consultants with experiences on similar domain or other domain
Technologies	• Tools or platforms of data acquisition and preprocessing • Tools or platforms of KG's representation, extraction, storage and query • Tools or platforms of crowdsourcing	• KG' databases and the suite of model and algorithms • Domain-specific information systems
Collaboration within KG's Make or Use	• Tasks of KG's Make are supported by a series of relevant actions on structure, actor and technology levels	• Tasks of KG's Use are supported by a series of relevant actions on structure, actor and technology levels
Artifacts involving KG	• Domain-specific KG is both the immediate concrete outcomes of the Make of KG and the inputs of the Use of KG. • Needs and IS of KG-based applications are separated into the Needs as both the inputs of the Make of KG and the outputs derived from the concrete applications as well as the IS as the outputs of the Use of KG.	
Collaborative Interaction between KG's Make and Use	• Interactions between KG's Make and Use are a process of cyclic shaping that contribute to quality and applicability of KG through three cycles. • Each cycle involves a collaborative action. The actions includes conceptual adaptation, technological experimentation and systematic actualization.	

3.2 Collaborative Mechanism in the Make of KG

KG-Make Tasks

We classify the KG-Make tasks into the expertise-dominant task and technology-dominant task. Specifically, *the expertise-dominant task* refers to that the experts, according to their expertise on KG and specific domain, make the KG mainly by a manual means. The success of expertise-dominant depends on the experts' knowledge

and experiences but is difficult to deal with a large amount of text, where supports of technological tools and platforms are necessary. By contrast, *the technology-dominant task* refers to that the KG is made automatically through information technology tools or platforms. The task is suitable for textual information extraction but is often lack of relevance and accuracy, where text annotation and labeling are necessary.

KG-Make Actions

A series of KG-Make actions should be taken on the structure, actor, and technology levels in order to improve depth, breath and quality of KG result (See Table 2). This study also combines both the tasks in practice and the bidirectional approach contribute to collaboration between KG's Make and Use (See 3.4).

Table 2. KG-Make actions on the structure, actor and technology levels

Structure level	Actor level	Technology level
• Constructing a technical-oriented group consisting of multi-role actors who can complement each other • Establishing a professional crowdsourcing community for annotation and labeling that are significant parts of KG's Make	• Setting up the role of KG architects owning the ambidextrous skill to coordinate the Make of KG • Accumulating domain-specific experts and technical experts into technical-oriented group and recruiting some experts and employees into crowdsourcing community through unified working criteria	• Acquiring the tools or platforms of data acquisition & preprocessing • KG's representation, reasoning & query and crowdsourcing through cloud computing environments or open-source software • Integrating the tools and platforms through enacting technological criteria

3.3 Collaborative Mechanism in the Use of KG

KG-Use Tasks

We classify the KG-Use tasks into the graph-visualization task and the graph-association task. Specifically, *the graph-visualization task* refers to basic applications of KG, such as retrieval and reasoning. The task also can be combined with concrete business context, but the business context is only based on the data visualization of KG, such as enhanced information index in knowledge service systems (e.g., [19]). By contrast, *the graph-association task* is defined as in-depth applications of KG embedded into concrete business processes or insights, such as personalized search and recommendation, context-based question answering and intelligence prediction (e.g., [20,21]).

KG-Use Actions

A series of KG-Use actions should be taken on the structure, actor and technology levels in order to explore thebusiness scenario of KG application and to develop KG-based IS (See Table 3). This study also aligns KG-Use tasks and KG-Make tasks in practice (See 3.4).

Table 3. KG-Make actions on the structure, actor, and technology levels

Structure level	Actor level	Technology level
• Constructing a product/ application-oriented group who are intended to use KG in IT/IS product/project • Aligning product/application group and IT/IS-oriented group through a digital innovation vision	• Improving KG-related knowledge of PM/BA and experts through a central education program • Recruiting KG consultants from university or professional institute	• Deploying the KG databases in the concrete IS environments • Integrating KG into domain-specific IS trough combination of models and algorithms

3.4 Collaborative Mechanism between KG's Make and Use

Interactive Action-1: Conceptual Adaptation

The uncertainty of development and utilization of emerging digital technologies [22] deeply affects processes of the Make and Use of KG. At the beginning of technological growth, especially the first two stages of Gartner Hype Cycle [23], the expectations of Makers and Users are different. The differences between KG's Make and Use reveal the intangible boundary [24] that originates from actors' cognitions. Thus, the conceptual adaptation is the first action that span the boundary.

On the one hand, most actors related to KG's Make and Use are different. However, the domain-specific experts should be assigned to participate in each tasks of KG. The experts are familiar with the business needs and responsible for conceptual definitions such as domain-specific entities and relations. The unified and unambiguous criteria of conceptual definitions are the common basis of KG's Make and Use. On the other hand, the Make of KG is subject to the Use of KG, the Use-ends groups may be too aggressive to cause conflicts among different technological groups. Thus, due to their ambidextrous capacities on both knowledge management and big data analytics, the Make-ends KG architects should be assigned to be the boundary spanners. The KG architects should coordinate technological groups at both ends to design the KG's criteria of representation, storage and query, so that the KG outcomes and KG-based IS needs can be in alignment.

Interactive Action-2: Technical Experimentation

Unlike mature technologies, KG as an emerging digital technology does not have any existing use cases in most domains [18]. Moreover, because of domain-specific features, a certain domain-specific KG and its application approach, to a large extent, cannot be reused directly in others, e.g., the use cases of entities and relations in the biomedical domain can be extracted from existing professional database such as Comparative Toxicogenomics Database [25], while management-science domain has no similar database that can offer us unified and unambiguous concepts. That means people should start "from zero to one", where a series of technical experimentations is necessary. The technical experimentation is an heuristic action that explores KG's Use scenarios and typical use cases which will be leveraged to guide a large-scale Make and Use of KG in subsequent IT/IS projects.

The process of technical experimentation should be bidirectional. On the one hand, KG architects and domain-specific experts, in a certain scope of business, work together to build some expertise-dominant instances by KG representation tools. The instances may inspire PM/BA to create use scenarios, e.g., when we build seven KG instances of IS-related theory, the teachers found that KG could be used in MOOC as a graph-visualization task that provided students with knowledge map and index. On the other hand, technology-dominant KG can be extracted and built from high-quality documents by KG extraction tools (e.g., OpenIE, Python or other tools). Meanwhile, according to need description of PM/BA, KG architects and technical experts will improve corresponding algorithms (e.g., syntactic dependency) to optimize KG results at the Make-ends and KG use case at the Use-ends. Furthermore, we found that two types of KG created by the approaches above could be combined again on the semantic level by NLP tools. The overlaps and relevant parts of KG could guide us to further adjust algorithms of KG's Make and needs of KG's Use.

Interactive Action-3: Systematic Actualization

On the basis of the technical experimentation, the systematic actualization is a large-scale action that realizes both the domain-specific KG and the KG-based IS. On the one hand, KG architects and technical experts, according to the results of experimentation, enact criteria for text annotation and labeling and design corresponding tasks and workflow. The employees should be selected or recruited from both domain-specific users and students and conduct the annotation and labeling by the crowdsourcing platform. KG technical experts, according to the results of annotation and labeling, further train the algorithms of KG's Make by machine learning. On the other hand, IS technical experts integrate the KG results from the Make-ends into concrete business system development. Whether the graph-visualization task or the graph-association task, KG architects and PM/BA should evaluate the application results by developing a set of indicators.

Obviously, the process of KG's Make and Use is endless and continuously improved. Thus, all the tasks of KG are application-oriented. The results of KG's Use may not only lead to the adjustment even reset of KG's Make but also contribute to

new needs of KG applications. Accordingly, the new needs may be fed back to *the technical experimentation* or to *the conceptual adaptation* due to the unexpected consequences.

4 Conclusion

So far, our KG project is being in the prototype run while we are proceeding in examining theory-evidence-findings alignment. Simultaneously, given both the limited space of this article and the research in progress, we only summarize the preliminary concepts of framework. In the next step, we will describe the details of our action design research and analysis, further refine the explanation to the findings, and derive the final design theory of KG's Make and Use.

Our preliminary findings may unveil *a collaborative-activity framework of KG's Make and Use* which address the research gap regarding the collaborative activities on the Make and Use of KG as an emerging technology. The framework combines the socio-technical factors, identifies two tasks of KG applications – Make and Use, and further reveals collaboration mechanisms regarding socio-technical actions not only within the Make and the Use separately but also between them through three interaction cycles.

Thus, our findings may make both important theoretical and practical contributions. First, this study contributes not only to the KG/AI and knowledge collaborative networks literature but also to the STS literature. Second, organizations can use our findings to develop their own domain-specific KG and KG-based IS, and by doing so, increase the success rate of their KG applications. Furthermore, we hope our findings will have impact on the future applications of KG in diverse domains.

Acknowledgement: Funding for this research was provided by "Research on scene behavior intelligence in collaborative network based on data analysis" (NSFC 61972029), "the Fundamental Research Funds for the Central Universities" (2019RCW011) and "Beijing Jiaotong University Education Foundation" (03060083).

References

1. Andreas, B., Helmut, N.: THE KNOWLEDGE GRAPH COOKBOOK RECIPES THAT WORK. edition mono/monochrom, Vienna (2020)
2. Yan, J.H., Wang, C.Y., Cheng, W.L., Gao, M., Zhou, A.Y.: A retrospective of knowledge graphs. Frontiers of Computer Science 12 (1):55-74 (2018)
3. Hao-Yu, D., Lan, W.: MOOC Instruction Optimization in "Emergency and First Aid"—From the Perspective of the Knowledge Graph. DEStech Transactions on Computer Science and Engineering (csma) (2017)
4. Xiao, Y.H.: Knowledge Graph and Cognitive Intelligence. Zhangjiang Technology Review (04):30-33 (2019). (in Chinese)

5. Chen, G.Q., Wu, G., Gu, Y.D., Lu, B.J., Wei, Q.: The challenges for big data driven research and applications in the context of managerial decision-making: Paradigm shift and research directions. JOU R NAL OF MANAGEMENT SCIENCES IN CHINA 21 (07):1-10 (2018). (in Chinese)
6. Camarinha-Matos, L., Afsarmanesh, H.: Classes of collaborative networks. In: IT Outsourcing: Concepts, Methodologies, Tools, and Applications. IGI Global, pp 364-370 (2010)
7. Chen, X.J., Jia, S.B., Xiang, Y.: A review: Knowledge reasoning over knowledge graph. Expert Systems with Applications 141:112948 (2020)
8. Cao, Q., Zhao, Y.M.: the realization process and application of knowledge map. Information Studies:Theory & Application 38 (12):127-132 (2015)
9. Lyytinen, K., Newman, M.: Explaining information systems change: a punctuated socio-technical change model. European Journal of Information Systems 17 (6):589-613 (2008)
10. Dremel, C., Herterich, M.M., Wulf, J., Vom Brocke, J.: Actualizing big data analytics affordances: A revelatory case study. Information & Management 57 (1):103121 (2020)
11. Ackoff, R.L.: From data to wisdom. Journal of applied systems analysis 16 (1):3-9 (1989)
12. Xiao, Y.H., Xu, B., Lin, X., Li, Z.X., Peng, P., Zheng, W.G., Shao, B., He, L., Yang, D.Q., Cui, W.Y.: Knowledge Graph： Concepts and Techniques. PUBLISHING HOUSE OF ELECTRONICS INDUSTRY, Beijing (2020). (in Chinese)
13. Miller, J.J.: Graph database applications and concepts with Neo4j. In: Proceedings of the Southern Association for Information Systems Conference, Atlanta, GA, USA, 2013. vol 36. (2013)
14. Linden, A., Fenn, J.: Understanding Gartner's hype cycles. Strategic Analysis Report Nº R-20-1971 Gartner, Inc:88 (2003)
15. Xu, X.F., Hao, J., Deng, Y.R., Wang, Y.: Design optimization of resource combination for collaborative logistics network under uncertainty. Applied Soft Computing 56:684-691 (2017)
16. Zhang, G.Q., Shang, J., Li, W.L.: Collaborative production planning of supply chain under price and demand uncertainty. European Journal of Operational Research 215 (3):590-603 (2011)
17. Wang, H.F., Chen, M.L.: Review for Information System Sociomateriality. Science and Technology Management Research (18):153-161 (2017). (in Chinese)
18. Du, W.Y., Pan, S.L., Leidner, D.E., Ying, W.C.: Affordances, experimentation and actualization of FinTech: A blockchain implementation study. The Journal of Strategic Information Systems 28 (1):50-65 (2019)
19. Dietz, L., Kotov, A., Meij, E.: Utilizing knowledge graphs for text-centric information retrieval. In: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, 2018. pp 1387-1390 (2018)
20. Kumar, A.J., Schmidt, C., Köhler, J.: A knowledge graph based speech interface for question answering systems. Speech Communication 92:1-12 (2017)
21. Zhang, Y.X., Yang, K.L., Du, W., Xu, W.: Predicting Stock Price Movement Direction with Enterprise Knowledge Graph. In: PACIS, 2018. p 237 (2018)
22. Roca, J.B., Vaishnav, P., Morgan, M.G., Mendonça, J., Fuchs, E.: When risks cannot be seen: Regulating uncertainty in emerging technologies. Research Policy 46 (7):1215-1233 (2017)
23. Kasey, P.: 5 Trends Appear on the Gartner Hype Cycle for Emerging Technologies, 2019. https://www.gartner.com/smarterwithgartner/5-trends-appear-on-the-gartner-hype-cycle-for-emerging-technologies-2019/.
24. Du, W.Y., Pan, S.L.: Boundary spanning by design: toward aligning boundary-spanning capacity and strategy in it outsourcing. IEEE Transactions on Engineering Management 60 (1):59-76 (2012)

25. Wang, Q.Y., Huang, L.F., Jiang, Z.Y., Knight, K., Ji, H., Bansal, M., Luan, Y.: Paperrobot: Incremental draft generation of scientific ideas. arXiv preprint arXiv:190507870 (2019)