IdeaNet: Potential Opportunity Discovery for Business Innovation

O-Joun Lee[1] , **Sung Youn Park**[2] and **Jin-Taek Kim**[1,*]

[1]Future IT Innovation Laboratory, Pohang University of Science and Technology, Republic of Korea
[2]Crevate, Republic of Korea
ojlee112358@postech.ac.kr, ceo1@crevate.com, jintaek@postech.ac.kr

Abstract

This paper introduces an ongoing project for mining potential business opportunities from the existing business innovation cases. Crevate (a consulting company) has collected thousands of cases from news articles and columns that are in textual narratives. We aim to transform the cases into knowledge graphs that cover what kinds of ideas (e.g., untact) are applied to which industrial areas (e.g., cafe). By using link prediction methods, we will be able to evaluate the prominence of combinations (e.g., untact cafe) between ideas and domains. This study focuses on explaining the collected cases and knowledge graphs with running examples.

1 Introduction

In the business intelligence area, data sources are frequently textual narratives. Crevate, a business consulting company, has collected thousands of business innovation cases in short textual narratives. Also, a few tools for supporting the potential idea discovery follows fixed scenarios, and its results are in narratives [Wang and Ohsawa, 2011]. Furthermore, a few studies [Segura *et al.*, 2018; Marjanovic, 2016] attempted to employ storytelling for helping business users explore visual analytics results. However, most of the existing studies focused on discovering bursty topics in social media texts [Yan *et al.*, 2015; Lee *et al.*, 2018; Lee *et al.*, 2017]. There have not been studies that consider narrative characteristics consistently from collecting existing business innovation cases, via discovering potential business opportunities, to providing the opportunities in explainable ways.

In this study, we introduce an ongoing project that aims to propose a representation model for the business innovation cases in the narrative form. Based on the representation, we focus on (i) analyzing narratives for business innovation cases, (ii) discovering the potential business opportunities, and (iii) representing the opportunities as narratives again.

2 Dataset

Crevate[1] has collected 3000 business innovation cases to inspire its staff and clients. The staff of Crevate have investigated the cases introduced by news articles. For example, 'a company applied bending machines on cafes.' They have made descriptions for each case that consist of 5 to 6 sentences by summarizing the news articles. Also, they have titled the cases (e.g., automated cafe), classified the cases according to merchandise types (e.g., physical products, services, etc.), and tagged industrial areas and keywords. This dataset includes short narratives with titles, summaries, categories, and tags. The staff have also composed a three-layered taxonomy of the innovation cases. The lowest layer includes categories, such as 'feeling tastes of the world,' 'feeling local tastes,' 'including numerous options,' and so on. The middle layer consists of 'experiencing the world,' 'experiencing the locality,' 'providing options to customers,' and so on. The highest layer contains 'extending experiences,' 'maximizing options,' and so on. Therefore, we can use the dataset for evaluating various tasks in the narrative analysis (e.g., title generation, summarization, tagging, classification, labeling, etc.). Each case has been annotated with 15 attributes, as presented in Table 1.

The cases consist of three parts: (i) what (approaches), (ii) where (domains or industrial areas), and (iii) how (specific methods). In the example of Table 1, 'Universal Yums[2]' applied the subscription and delivery services on snacks by providing diverse local tastes. As a similar example, 'Mouth[3]' applied the subscription and delivery on indie foods by providing local foods from all over America on a website.

As shown in the summary and description in Table 1, these cases are simple and formalized narratives. Also, the title, industry, and tags provide us guidelines for analyzing the narratives. With Part-Of-Speech (POS) tagging, we can find the three parts from sentences in the summary and description. Objects of sentences will be related to domains of the cases (e.g., a box of snacks). From verbs, we can find the

*Corresponding Author.

In: A. Jorge, R. Campos, A. Jatowt, A. Aizawa (eds.): Proceedings of the first AI4Narratives Workshop, Yokohama, Japan, January 2021, published at http://ceur-ws.org

[1]https://crevate.com/
[2]https://www.universalyums.com/
[3]https://www.mouth.com/

Attribute	Description	Example
Title	The name of services/products	Snack boxes with tastes of the world
Summary	Abstract of descriptions for services/products	This service monthly delivers a box of snacks which are representative of a country.
Description	Content and characteristics of services/products	Universal Yums sends its staff to on-spot-surveys for finding local snacks. The staff tastes and scores the local snacks to compose monthly snack boxes. Scoring criteria are distinct local characteristics, diversity of tastes, and so on.
Type	Kinds of business innovations in seven categories (e.g., products, services, etc.)	Service
Industry	Industrial areas of services/products	Food and beverage
Year	The year of release	2014
Nation	Service regions	United States of America
Company	The name of company	Universal Yums
Image	Sample images	–
Video	Promotion videos	–
Link	Promotion websites	https://www.universalyums.com/
Tags	Keywords related to the innovation cases	Subscription; Delivery; Snacks
Category 1	Categories on the lowest layer in the taxonomy	Feeling tastes of the world
Category 2	Categories on the middle layer	Experiencing the world
Category 3	Categories on the highest layer	Extending experiences

Table 1: Descriptions and examples of the business innovation dataset.

approaches employed in the cases (e.g., delivers). Lastly, adverbs will reflect the specific methods for applying the approaches (e.g., monthly). WordNet can also help us find the three components from the descriptions. For example, since 'snack' is a hyponym of 'Food and beverage,' we can assume that it will indicate the domain of the case. In further research, we will also consider an automated extension of the dataset by crawling new articles.

3 Business Innovation Knowledge Graph

By extracting approaches (what), domains (where), and methods (how) from the business innovation cases, we can model relationships between the three components. As shown in the two examples, lots of cases among the dataset share the approaches and domains, even sometimes a part of the methods. If we represent relationships between components in a case, we will be able to search the cases more semantically than before. Also, representing the relationships between cases will enable us to examine correlations between the approaches, domains, and methods (e.g., synergies and trade-offs). These relationships can be modeled as a knowledge graph. Fig. 1 presents the knowledge graph-based representations for 'Universal Yum' and 'Mouth.'

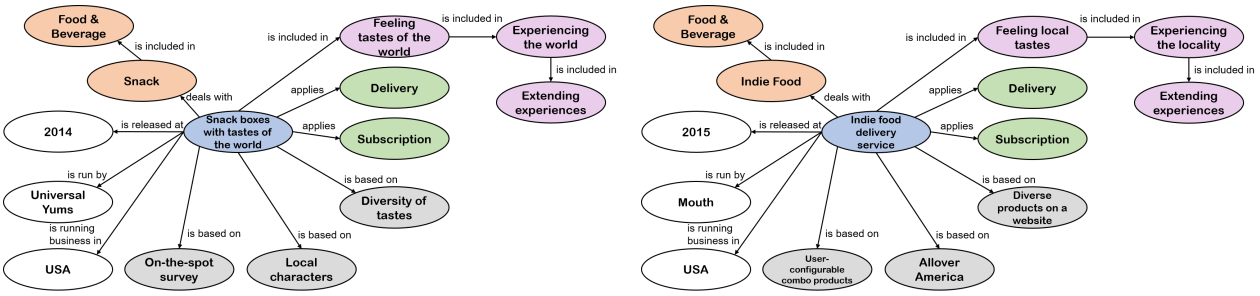
We define the representation model that has 6 kinds of nodes and 7 sorts of relations. The types of nodes are as follows.

- Case: This node indicates the business innovation case itself. The other types of nodes are to describe the properties of this node.
- Domain: This node presents industrial areas of the case. Domains can be described in multiple layers (e.g., Snacks $\in$ Food & Beverage). Also, a case can have multiple domains (e.g., 'functional foods' and 'cosmetics').
- Approach: We also annotate which approaches are used in the cases. As with the above, one case can apply multiple approaches (e.g., 'delivery' and 'subscription').
- Method: Methods means specific strategies that are used in the cases. 'Universal Yum' has conducted 'on-the-spot surveys' for finding snacks that have 'local characters' to satisfy 'diversity of tastes.'
- Category: This node presents the three-layered taxonomy of the innovation cases.
- Metadata: Excluding the semantic information, the cases include additional information, such as company names, release years, service regions, and so on.

The types of relations are as follows.

- 'deals with' connects cases and domains of the cases.
- 'applies' links cases and their approaches.
- 'is based on' annotates methods of the cases.
- 'is included in' links subclasses and superclasses.
- 'is run by,' 'is running business in,' and 'is released at' commonly indicate additional information of the cases.

By representing the cases in a single knowledge graph, we can examine which approaches and domains frequently appear together, as displayed in Fig. 2. On the relations, we can assign weights based on frequency of the cases. Chen et al. [2019a] also used a network model for idea mining. However, their model only combines two concepts (e.g., 'leaf' and



(a) Snack boxes of 'Universal Yum'

(b) Indie food market of 'Mouth'

Figure 1: Knowledge graphs for describing individual business innovation cases. Colors indicate kinds of nodes. Blue nodes are the innovation cases, orange ones denote domains of cases, green ones are their approaches, gray ones describe their specific methods, pink ones present categories of the cases, and white nodes include meta information.

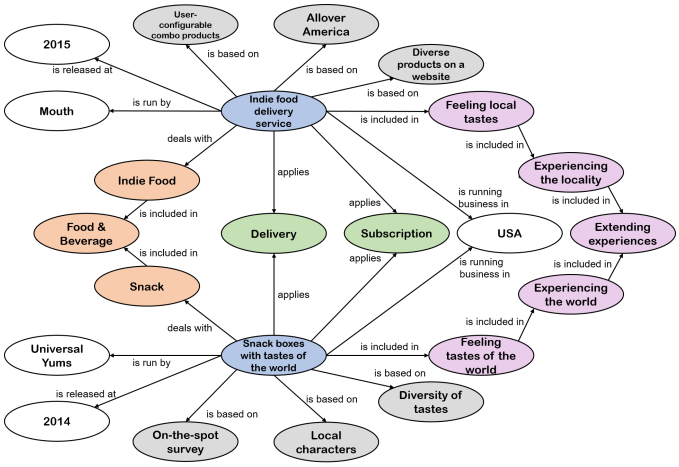


Figure 2: A unified knowledge graph for describing correlations between the approaches, domains, and methods in business innovation.

'spoon') to discover novelty (e.g., 'leaf-shaped spoon'). On the other hand, we depict components of innovation cases as detail as possible.

4 Potential Opportunity Discovery

Crevate has merely provided a search engine for the existing innovation cases by using the approaches and domains. This retrieval has been conducted by a three-step procedure. Users insert business domains that they want. The search engine shows a list of candidate approaches to the users. Finally, after the users choose one of the approaches, the search engine provides innovations cases that have conducted on the selected domain with the chosen approach. This service can provide only existing cases.

Nevertheless, by analyzing the knowledge graph, we can examine the potentials of the cases, whether they have been tried or remained as terra incognita. There have been various studies [Rossi *et al.*, 2020] for predicting links in knowledge graphs based on: tensor factorization [Trouillon *et al.*, 2017], network embedding [Chen *et al.*, 2019b; Kazemi and Poole, 2018], and affinity propagation [Wang *et al.*, 2018]. To discover the potential opportunities, we have to consider both affinity between nodes and structures of the knowledge graph. Let suppose that 'cafe' and 'bookstore' are commonly connected to 'providing cozy spaces' with high affinity. If 'cafe' and 'untact services' also have high affinity, we can expect that 'untact services' is applicable to 'bookstore,' as well. Therefore, we will apply metapath-based network embedding methods (e.g., HIN2Vec [Fu *et al.*, 2017]), which are effective for representing both affinity and network structures. Based on the prediction results, we can discover more potential combinations from all possible combinations. Additionally, we also plan to generate descriptions of the discovered potential opportunities in the narrative form.

5 Conclusion

This paper presents backgrounds and approaches of the ongoing project. In further research, we will focus on clarifying our methodologies considering characteristics of the dataset.

Acknowledgments

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICT Consilience Creative program (IITP-2019-2011-1-00783) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation).

References

[Chen *et al.*, 2019a] Liuqing Chen, Pan Wang, Hao Dong, Feng Shi, Ji Han, Yike Guo, Peter R.N. Childs, Jun Xiao, and Chao Wu. An artificial intelligence based data-driven approach for design ideation. *Journal of Visual Communication and Image Representation*, 61:10–22, May 2019.

[Chen *et al.*, 2019b] Mingyang Chen, Wen Zhang, Wei Zhang, Qiang Chen, and Huajun Chen. Meta relational learning for few-shot link prediction in knowledge graphs. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP 2019)*, pages 4216–4225, Hong Kong, China, November 2019. Association for Computational Linguistics.

[Fu *et al.*, 2017] Tao-Yang Fu, Wang-Chien Lee, and Zhen Lei. HIN2vec: Explore meta-paths in heterogeneous information networks for representation learning. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (CIKM 2017)*, pages 1797–1806, Singapore, November 2017. ACM Press.

[Kazemi and Poole, 2018] Seyed Mehran Kazemi and David Poole. Simple embedding for link prediction in knowledge graphs. In *Advances in Neural Information Processing Systems 31 - Proceedings of the 32nd Annual Conference on Neural Information Processing Systems (NeurIPS 2018)*, pages 4289–4300, Montréal, Canada, December 2018.

[Lee *et al.*, 2017] O-Joun Lee, Hoang Long Nguyen, Jai E. Jung, Tai-Won Um, and Hyun-Woo Lee. Towards ontological approach on trust-aware ambient services. *IEEE Access*, 5:1589–1599, February 2017.

[Lee *et al.*, 2018] O-Joun Lee, Yunhu Kim, Long Nguyen Hoang, and Jai E. Jung. Multi-scaled spatial analytics on discovering latent social events for smart urban services. *Journal of Universal Computer Science*, 24(3):322–337, March 2018.

[Marjanovic, 2016] Olivera Marjanovic. Empowering business users to explore visual data through boundary objects and storytelling. In *Proceedings of the 49th Hawaii International Conference on System Sciences (HICSS 2016)*, pages 5032–5041, Koloa, HI, USA, January 2016. IEEE Computer Society.

[Rossi *et al.*, 2020] Andrea Rossi, Donatella Firmani, Antonio Matinata, Paolo Merialdo, and Denilson Barbosa. Knowledge graph embedding for link prediction: A comparative analysis. arXiv:2002.00819, March 2020.

[Segura *et al.*, 2018] Vinicius Segura, Juliana Jansen Ferreira, and Simone D. J. Barbosa. BONNIE: Building online narratives from noteworthy interaction events. In *Proceedings of the 2018 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC 2018)*, pages 233–238, Lisbon, Portugal, October 2018. IEEE Computer Society.

[Trouillon *et al.*, 2017] Théo Trouillon, Christopher R. Dance, Éric Gaussier, Johannes Welbl, Sebastian Riedel, and Guillaume Bouchard. Knowledge graph completion via complex tensor factorization. *Journal of Machine Learning Research*, 18:130:1–130:38, November 2017.

[Wang and Ohsawa, 2011] Hao Wang and Yukio Ohsawa. iChance: A web-based innovation support system for business intelligence. *International Journal of Organizational and Collective Intelligence*, 2(4):48–61, October 2011.

[Wang *et al.*, 2018] Hongwei Wang, Fuzheng Zhang, Jialin Wang, Miao Zhao, Wenjie Li, Xing Xie, and Minyi Guo. RippleNet: Propagating user preferences on the knowledge graph for recommender systems. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management - (CIKM 2018)*, pages 417–426, Torino, Italy, October 2018. ACM Press.

[Yan *et al.*, 2015] Xiaohui Yan, Jiafeng Guo, Yanyan Lan, Jun Xu, and Xueqi Cheng. A probabilistic model for bursty topic discovery in microblogs. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI 2015)*, pages 353–359, Austin, Texas, USA, January 2015. AAAI Press.