Technological Forecasting & Social Change xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

Toward data-driven idea generation: Application of Wikipedia to morphological analysis☆

Heeyeul Kwon[a], Yongtae Park[a], Youngjung Geum[b,*]

[a] Department of Industrial Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul, Republic of Korea
[b] Department of Industrial & Information Systems Engineering, Seoul National University of Science and Technology, 172 Gongreung 2-dong, Nowon-gu, Seoul, Republic of Korea

ARTICLE INFO

Keywords:
Idea generation
Ideation
Morphological analysis
Wikipedia
Big data

ABSTRACT

The generation of new and creative ideas is vital to stimulating innovation. Morphological analysis is one appropriate method given its objective, impersonal, and systematic nature. However, how to build a morphological matrix is a critical problem, especially in the big data era. This research focuses on Wikipedia's case-specific characteristics and well-coordinated knowledge structure and attempts to integrate the platform with morphological analysis. In details, several methodological options are explored to implement Wikipedia data into morphological analysis. We then propose a Wikipedia-based approach to the development of morphological matrix, which incorporates the data on *table of contents*, *hyperlinks*, and *categories*. Its feasibility was demonstrated through a case study of drone technology, and its validity and effectiveness was shown based on a comparative analysis with a conventional discussion-based approach. The methodology is expected to be served as an essential supporting tool for generating creative ideas that could spark innovation.

1. Introduction

Innovation has been a central issue in both academia and practice. The general definition of innovation is discussed in numerous research studies. Evangelista et al. (1998) defined innovation as a process from research to invention and then a diffusion of a new technique. A study by Drucker (1985) defined as "the act that endows resources with a new capacity to create wealth". In recent years, numerous scholars emphasized the importance of generating creative ideas in achieving such an innovation (Girotra et al., 2010; Rietzschel et al., 2014) and further proposed the notion that creative and innovative ideas originate from the right combination of disparate bodies of existing knowledge (Björk and Magnusson, 2009; Dosi, 1982; Geschka, 1983; Nakamura et al., 2015; Schilling and Green, 2011; Ward, 2004).

Morphological analysis has been particularly employed as a prominent tool for generating new ideas. It is a method that leads to structured inventions by determining all possible alternatives for solving a certain problem (Wissema, 1976; Yoon and Park, 2005; Geum et al., 2016). This technique has two strong advantages for idea generation. First, morphological analysis decomposes a complex system into parts and systematically rearranges combinations to generate ideas. This can be considered a "combinative" characteristic of the technique that sparks an abundant supply of creative ideas. Second, morphological analysis breaks down the target subject and reconstructs them to explore unprecedented structures. This can be considered an "inventive" characteristic of the technique that explores many different novel ideas, thus providing possible solution options for a given context.

Despite its popularity, one innate shortcoming of morphological analysis must be solved: the involvement of human subjectivity. The entire process of decomposing and restructuring into new knowledge can exist in the first place only if the morphological matrix is designed properly and specifically. This preprocessing step is the most fundamental step when considering the notion of GIGO–short for garbage in, garbage out. Previous studies, however, handled the morphology building process in a qualitative manner that depended on experts' opinions. Regardless of their advantages in capturing high reliability and validity, the involvement of human judgment was subject to numerous cognitive biases that could subsequently lead to a decrease of team performance and serious misjudgments (Zec et al., 2015). Furthermore, the conventional participatory process can no longer be applied to recently emerged products or services, which are so complex

☆ It is confirmed that this item has not been published elsewhere.

* Corresponding author at: Department of Industrial & Information Systems Engineering, Seoul National University of Science and Technology, 172 Gongreung 2-dong, Nowon-gu, Seoul 139-746, Republic of Korea.
E-mail address: yjgeum@seoultech.ac.kr (Y. Geum).

https://doi.org/10.1016/j.techfore.2018.01.009
Received 19 January 2017; Received in revised form 3 January 2018; Accepted 9 January 2018

that a concrete decomposition cannot be achieved solely by a handful of professionals. For these reasons, there is in need of a shift toward a data-driven methodology, which could offer more objective and automatic results.

Several attempts have been made to suggest a data-driven approach of conducting morphological analysis, and patent data was generally incorporated as the main data source (Lee et al., 2007; Yoon et al., 2008; Yoon et al., 2014). However, one of the major limitations was that they were restricted to a single knowledge source of patent. Although it is renowned for its high reliability and usefulness, patent data often could not ensure up-to-dated, non-technological, or case-specific idea generation. To address this limitation, Geum and Park (2016) suggested WordNet-based morphological analysis. Their study highlighted the hierarchical network structure of WordNet, which has the capability of analytic and objective morphology development. However, WordNet was still subject to data rigidity and lacking domain-specific knowledge.

In response, this paper proposes a novel method of data-driven morphology building process using Wikipedia data. Wikipedia is a massive online repository of collective intelligence. The utilization of Wikipedia can be an excellent way to address the rigidity of WordNet. In terms of practicability, Wikipedia contains information from a wide range of fields and expands the data set to a new boundary. Unlike WordNet, Wikipedia offers a significant amount of case-specific knowledge that could better stimulate more specialized and practical idea generation. In terms of flexibility, Wikipedia provides hierarchical relationships of concepts, which enable a coherent matrix development of morphological analysis.

One thing to note is that the objective of this research is neither to propose the most effective method of developing creative ideas nor to supersede conventional idea generation techniques. The approach is rather a heuristic process model, attempted to demonstrate the key role of Wikipedia in morphology building step for ensuring conceptual breadth and facilitating creativity. The remainder of this research is organized as follows. Section 2 illustrates how the research of creative idea generation and morphological analysis has been previously carried out. Then, Section 3 describes the types of Wikipedia information to apply in the development of a morphological matrix and offers several alternatives to achieving this objective. With this background, we propose a research framework of constructing a Wikipedia-based morphological matrix in Section 4. Section 5 demonstrates the feasibility of our approach through the illustrative case study of drone technology. Section 6 further shows its validity and effectiveness based on a comparative analysis with a conventional discussion-based approach. Lastly, Section 7 discusses the intrinsic shortfalls of applying Wikipedia data in morphological analysis.

2. Related studies

2.1. Generating creative ideas

Generating new ideas has been a constant concern. A considerable literature has, thus, endeavored to address novel methods of idea generation. Starting with Osborn's brainstorming (Osborn, 1957), diverse idea generation techniques came into existence, including brainwriting, checklists, and synectics (Geschka, 1983; Ivanov and Cyr, 2014; VanGundy, 1981). They all made a tacit assumption: *quantity breeds quality* (Rietzschel et al., 2014). In other words, these techniques have assumed that as more ideas are created, the greater the possibility that creative ideas are found among them (Diehl and Stroebe, 1987; Girotra et al., 2010). As suggested by Shah et al. (2003), idea generation methodologies are classified into two primary groups: intuitive and logical. Intuitive methods are divided into germinal, transformational, SMAPER, progressive, C-sketch, organizational, fishbone diagram, and hybrid. Logical methodologies are divided into history-based methods and analytical methods such as TRIZ.

However, in recent years, the quality of the ideas has been questioned (Girotra et al., 2010; Rietzschel et al., 2014). Most studies underlined the concept of creativity as improving the quality (Rietzschel et al., 2014) and began to facilitate this improvement in idea generation process. Most studies demonstrated in a qualitative manner. For instance, Girotra et al. (2010) proposed the hybrid structure-in which individuals first work independently and then work together -and identified the structure's superior performance in terms of both quantity and quality. Rietzschel et al. (2014) applied a new brainstorming experiment and demonstrated that ideas became more creative as problems narrowed and instructions became more creative. In contrast, several studies attempted to demonstrate the improvement of quality in a quantitative manner. If a quantitative approach is taken, what is the best scale that determines a creative idea? According to Simonton (2013), the simplest but most useful answer is to treat criteria as dichotomous features. For example, a creative idea is original, rather than unoriginal, and useful, rather than useless. However, a definitional disagreement regarding the concept of creativity is inevitable given its subjective and versatile characteristics. Some research considered that creative ideas are both original and feasible (Diehl and Stroebe, 1987; Rietzschel et al., 2014; Sternberg, 1985). Other studies highlighted the concept of interventions during the generation process and demonstrated that ideas derived in the face of greater obstacles are considered more creative (Ward, 2004). Boden (2004) offered three criteria of novel, valuable, and surprising. As shown, the subjective concept of creativity was described through other subjective notions, such as original, novel, and valuable.

More recent studies focused on the methodology itself that generates new and creative ideas, and many of those stressed the underlying concept that creative ideas are crafted from existing knowledge by identifying novel combinations of previously separated ideas or concepts (Björk and Magnusson, 2009; Dosi, 1982; Geschka, 1983; Nakamura et al., 2015; Schilling and Green, 2011; Ward, 2004). In other words, a method that can systematically combine valuable knowledge from novel data sources is expected to generate creative ideas.

2.2. Data-driven morphological analysis

Morphological analysis is the study of basic forms or patterns of a thing (Ritchey, 2011). When viewed from the perspective of *totality of things* (Zwicky, 1957), a certain object can be divided into multiple parts, and different arrangements of those parts could conform to create different wholes. The decomposition can be organized into a table, so-called morphological matrix, composed of dimensions and values. A conventional example of a morphological matrix is shown in Table 1.

As an illustration, an unmanned aerial vehicle (UAV)-commonly known as drone-can be decomposed into four dimensions: body, sensors, actuators, and software. Generally, the body design of small UAV has 4 propelled rotors and is called quadcopter. If the body changes from 4 to 8 rotors, the UAV is an octocopter and serves a very different purpose than that of a quadcopter. Apparently, every dimension includes more than one value; for example, the body dimension comprises four, six, and eight rotor designs. When considering that a UAV is a

Table 1
Morphological matrix example for textile wet-processing system (Jones, 1976).

Textile additive	Fibre substrate	Process medium	Additive fibre mechanism
Dyestuff	Protein	Aqueous	Substantive
O.B.A.	Cellulose	Aqueous solution	Diffusion
Finish	Nylon	Organic solvent	Precipitation
Antistat	Polyester	Air gas vacuum	Reactive