The crowdsourcing data for innovation: Does it matter?

Ilmi, Zainal; Wijaya, Adi; Kasuma, Jati; Darma, Dio Caisar

European Journal of Management Issues

Volume 28(1-2), 2020, pp.3-12

DOI: 10.15421/192001

Received: 20 January 2020; 01 April 2020
Revised: 22 March 2020; 13 April 2020
Accepted: 18 June 2020
Published: 25 June 2020

UDC classification: 338

JEL Classification: O30, M15

The crowdsourcing data for innovation: Does it matter?

***Z. Ilmi*‡, *A. Wijaya*‡‡, *J. Kasuma*‡‡‡, *D. C. Darma*‡‡‡‡**

Purpose – to explore a crowdsourcing data-driven approach to construct crowdknowledge databases for innovation through supporting creative idea generation. In the approach, social media will be used as platforms to crowdsource knowledge for producing the databases.

Findings. Creativity is an essential element of innovation, but producing creative ideas is often challenging in design. Many computational tools have become available recently to support designers in producing creative ideas that are new to individuals. As a standard feature, most of the tools rely on the databases employed, such as ConceptNet and the US Patent Database. This study highlighted that the limitations of these databases have constrained the capabilities of the tools and, thereby, new computational databases supporting the generation of new ideas to a crowd or even history are needed. Crowdsourcing outsources tasks conventionally performed in-house to a crowd and uses external knowledge to solve problems and democratize innovation. Social media is often employed in crowdsourcing for a crowd to create and share knowledge.

Originality/Value. This paper proposes a novel approach employing social media to crowdsource knowledge from a crowd for constructing crowd knowledge databases.

Practical implications. The crowd knowledge database is expected to be used by the current computational tools to support designers producing highly creative ideas that are new to the crowd, in new product design, and ultimately to innovation.

Research limitations/Future research. In this study to provide insights and potential directions for future research are discussed that challenges of employing described approach.

Paper type – theoretical.

Keywords: creativity; project management; data-driven design; innovation; social media; IT-management.

‡*Zainal Ilmi,*
Senior Lecturer, Dr, Department of Management,
Faculty of Economics and Business,
Mulawarman Univesity, Samarinda, Indonesia,
e-mail: zainal.ilmi@feb.unmul.ac.id

‡‡*Adi Wijaya,*
Senior Lecturer, Department of Economics,
Faculty of Economics and Business,
Mulawarman Univesity, Samarinda,Indonesia,
e-mail: adi.wijaya@feb.unmul.ac.id,
https://orcid.org/0000-0003-1396-5891

‡‡‡*Jati Kasuma,*
Assoc. Prof., Faculty of Business and Management,
Universiti Teknologi Mara, Shah Alam, Malaysia,
e-mail: jati@uitm.edu.my,
https://orcid.org/0000-0002-4041-5468

‡‡‡‡*Dio Caisar Darma,*
Researcher, Department of Management,
Sekolah Tinggi Ilmu Ekonomi, Samarinda, Indonesia,
e-mail: diocaisar@stiesam.ac.id,
https://orcid.org/0000-0002-3287-7670

Reference to this paper should be made as follows:
Ilmi, Z., Wijaya, A., Kasuma, J., & Darma, D. C. (2020). The crowdsourcing data for innovation: Does it matter? *European Journal of Management Issues*, 28(1-2), 3-12. doi:*10.15421/192001*.

Дані краудсорсингу для інновацій: чи це має значення?

Зайнал Ільмі[‡], Аді Віджая[‡], Джаті Касума[‡‡], Діо Кайсар Дарма[‡‡‡]

[‡]Університет Мулаварман, Самарінда, Індонезія,
[‡‡]Universiti Teknologi Mara, Шах Алам, Малайзія,
[‡‡‡]Економічний коледж, Самарінда, Індонезія

Мета роботи – вивчити краудсорсинговий підхід, заснований на даних для створення баз даних краудсорсингу для інновацій за допомогою підтримки генерації творчих ідей. У цьому підході соціальні медіа будуть використовуватися як платформи для краудсорсингу знань для створення баз даних.

Результати дослідження. Творчість – важливий елемент інновацій, проте створення креативних ідей часто складне завдання у проектуванні. Наголошено, що нещодавно багато обчислювальних інструментів стали доступними для підтримки дизайнерів у виробленні нових творчих ідей для окремих замовників. Зазвичай більшість дизайнерських інструментів покладаються на використовувані бази даних, такі як ConceptNet та База даних патентів США. У цьому дослідженні підкреслено, що обмеження цих баз даних зменшують можливості інструментів, а отже, необхідні нові обчислювальні бази даних, що підтримують генерування нових ідей для краудсорсингу або навіть історії процесів. Під час краудсорсингу передають на аутсорсинг завдання, які зазвичай виконують власноруч, з метою використати зовнішні знання для вирішення проблем та демократизації інновацій. Наголошено, що соціальні медіа застосовують у краудсорсингу для накопичення та зовнішнього обміну знаннями.

Оригінальність/Цінність/Наукова новизна дослідження. Запропоновано новий підхід із використання соціальних медіа задля накопичення крауд-баз даних для інновацій.

Практичне значення дослідження. Очікується, що крауд-база даних знань може застосовуватися сучасними обчислювальними інструментами для підтримки дизайнерів, що виробляють висококреативні ідеї, які є новими для краудсорсингу, під час розробці нових продуктів і, зрештою, для інновацій.

Обмеження досліджень/Перспективи подальших досліджень. У цьому дослідженні обговорено проблеми застосування описаного підходу задля надання розуміння та потенційних напрямків для подальших досліджень.

Тип статті – теоретичний.

Ключові слова: творчість; проектне управління; інформаційне проектування; інновації; соціальні засоби комунікації; IT-менеджмент.

Данные краудсорсинга для инноваций: имеет ли это значение?

Зайнал Ильми[‡], Ади Виджая[‡], Джати Касума[‡‡], Дио Кейсар Дарма[‡‡‡]

[‡]Университет Мулаварман, Самаринда, Индонезия,
[‡] Universiti Teknologi Mara, Шах Алам, Малайзия
[‡‡‡]Экономический колледж, Самаринда, Индонезия

Цель работы – изучить краудсорсинговый подход, основанный на данных для создания крауд-баз данных для инноваций посредством поддержки генерации творческих идей. В этом подходе социальные медиа будут использоваться как платформы для краудсорсинга знаний для создания баз данных.

Результаты исследования. Креативность - важный элемент новаторства, но создание креативных идей часто является сложной задачей в проектировании. Отмечено, что в последнее время стало доступно множество вычислительных инструментов для поддержки проектантов в разработке новых творческих идей для отдельных заказчиков. В качестве стандартной функции большинства инструментов проектанты используют базы данных, такие как ConceptNet и Патентная база данных США. В исследовании показано, что ограничения этих баз данных урезают возможности инструментов и, следовательно, необходимы новые вычислительные базы данных, поддерживающие генерацию новых идей для краудсорсинга или даже для истории процессов. При краудсорсинге на аутсорсинг передаются задачи, обычно выполняемые собственными силами, с целью использовать крауд-знания в решении проблем и демократизации инноваций. Отмечено, что социальные сети часто используют в краудсорсинге для накопления и внешнего обмена знаниями.

Оригинальность/Ценность/Научная новизна исследования. Предложен новый подход использования социальных сетей при краудсорсинге для создания крауд-баз данных.

Практическое значение исследования. Ожидается, что база данных коллективных знаний будет использоваться текущими вычислительными инструментами для поддержки проектантов, создающих очень креативные идеи, которые являются новыми в краудсорсинге, в проектировании новых продуктов и, в конечном итоге, в инновациях.

Ограничения исследований /Перспективы будущих исследований. В этом исследовании обсуждаются проблемы использования описанного подхода для предоставления понимания и потенциальных направлений для будущих исследований.

Тип статьи – теоретический.

Ключевые слова: творчество; проектное управление; информационное проектирование; социальные медиа; IT-менеджмент.

1. Introduction

People in some countries use card payments less often than people in other countries and often prefer cash (*Yohannes, 2015*). The first proposed explanation for using cash instead of card for payment is control over money (*Kalckreuth et al., 2014*). One can assume that people in these countries – such as Japan, Germany and Spain – implicitly associate card payments with less control. In order to examine this assumption we decided to run the Implicit Association Test (IAT) (*Greenwald et al., 1998; Greenwald et al., 2002*), which is a reliable method for uncovering implicit associations between concepts on a deep psychological level.

A recent study by *Dasgupta and colleagues* (*2009*) shows that the implicit association can be influenced by the affective state participants are currently in. Consequently, we wanted to investigate whether the implicit association of cash and card payments with high or low control is altered in the specific affective state. In this manuscript, we try to provide an answer to the **research questions** "Do Germans implicitly associate the type of payment with different degrees of control?" and "Is the association influenced by the experienced affective state?"

We chose to perform our analysis in Germany. With this study, we contribute, first, to literature on human behaviour and IT-driven systems interaction. We investigate how implicit associations might shift human preferences concerning the type of payment. Second, we show that the German participants differ in their associations from other participants of the study. This is an interesting insight for cross-cultural research as well as for research on the use of cash payment. Our study uses an innovative method and opens avenues for researchers who seek to understand the effects of implicit association, prejudices, and perceptual biases on human-technology interaction.

2. Theoretical background

2.1. Relevance of payment processes

Payments are needed for all kinds of economic activity. Hence, companies as well as all other organizations have to implement payment processes. Currently, the payments industry is in a state of huge upheaval triggered by regulatory as well as political initiatives. These include the creation of the Single Euro Payments Area (SEPA), the establishment of instant payments, which is already on the way, the revised Payment Services Directive (PSD2), which became fully effective in 2019 in all EU member states, and the regulation on interchange fees (EU 2015/751). Most of the current projects serve the goal to harmonize the euro payments market in Europe, as well as to encourage more competition and open the market to new entrants.

Payments represent a major source of revenue for financial institutions. In fact, payments are not only a source of revenues, but they are the anchor product for various other services. In addition, payment information is a source of knowledge about data on customers, and an opportunity to generate points of reference into the processes of bank' customers – whether private, business, or institutional. Thus, losing stakes in payment transactions to other players would have disastrous consequences for banks.

Payment processes are provided mainly by banks and credit card organizations. However, the emergence of smartphones has allowed new players, such as large Internet and telecommunication enterprises, entering the market (PayPal, Apple, Facebook, Tencent, Alibaba to name a few). Furthermore, numerous companies from the fintech sphere (start-up companies in the financial services sector relying heavily on IT) have appeared on the payments market. The new players aim to integrate their payment services into the customers' processes, thereby capturing customer data, and tying the customers to the company.

2.2. Cash versus card payments

Why do people in countries such as Japan, Germany and Spain prefer cash over card payments? Often the explanation is control over money (*Kalckreuth et al., 2014*). Indeed, cash as well as debit cards are often seen as a monitoring and budgeting tool, especially in times of crises (*Hernandez et al., 2017*). The scholars argue that a substitution of cash by cards may slow down due to environmental turbulences. Therefore, electronic means of payment seem to be far from achieving the expected benefits of cash with regard to perceived control over own budget.

The appetite for cash seems to remain constant since people see cash not only as a mean of payment but mainly as a mean of value storage (*Bech et al., 2018*). Nevertheless, than type of payment has impact on the way consumers behave (*Runnemark et al., 2015*). For instance, *Falk and colleagues* (*2016*) found that the willingness-to-pay increases if consumers switch from cash to card or mobile payments. "Cash payments, which are more transparent than debit card transactions, make it easier to control spending and this effect is not solely due to cash-on-hand constraints" (*Runnemark et al., 2015, p. 286*). Therefore, it is necessary to understand the depth of the association between means of payment and the degree of control by an individual.

3. Methodology

3.1. Implicit Association Test

In an IAT, the participants are confronted with a series of stimuli which they have to sort. The sorting tasks are changing during the test. The regular IAT runs five trials during which the participants (1) have to sort words or pictures from one category (initial target concept discrimination, e.g., payment type: cash versus card), and (2) from the second category (associated attribute discrimination, e.g., control level: low versus high). Afterwards (3), the participants receive a task to make a combined sorting: if they see a picture with a cash payment or a word associated with high control, they have to sort it to the left; and if they see a picture with a card payment or a word associated with low control, then to the right (initial combined task, congruent condition). In the fourth trial (4), the participants have to perform a simple sorting of pictures associated with card or cash payment but the direction of sorting changes (reversed target concept discrimination, erasing of habits developed in the first trial). Finally (5), the participants have to sort words and pictures associated with cash payment and low control to the left and card payment and high control to the right (reversed combined task, incongruent condition).

In order to know the implicit association, the researcher has to calculate the mean time difference between trial five and trial three. A positive number would indicate that the congruent condition holds, i.e., the participants need less time to associate cash payment with high control and card payment with low control vice versa. A negative number indicates that the incongruent condition holds. Put differently, if a participant needs less time for sorting certain categories, we can assume that s/he implicitly associates these categories.

To conduct the test, our research team had to develop a set of stimuli for each category. If in the event of card or cash payment the pictures obviously belong to the specific category (*Fig. 1*), respective words needed to be found and pretested. We followed the procedure suggested by *Bogodistov and Dost* (*2017*).

1. Introduction

Creativity is connected to innovation via design (Han et al., 2018a), while creativity is often associated with idea generation. Idea generation (ideation) is the process of coming up with ideas during the early phases of design. It has been considered the foundation of innovation (*Sarkar & Chakrabarti, 2011; Cash & Štorga, 2015; Zainurossalamia et al., 2020*), which is also a significant element in business success (*Howard et al., 2011*). Therefore, generating creative ideas is essential for achieving innovation. However, it is always challenging for individuals to produce creative ideas due to limited knowledge, many existing ideas, time pressure, and a lack of creative mind (*Han et al., 2018a*). Knowledge is a significant resource in supporting innovation (*Bertola & Teixeira, 2003*), but it is difficult and time-consuming to collect information and knowledge for assisting idea generation. *Ullman* (*2010*) indicated that design engineers spend 60% of the time during the design process to explore the information and knowledge needed. Therefore, the designers need support in creating and leading to innovation, relevant knowledge, or a database containing the necessary knowledge.

There is an ever-growing interest in applying computational tools for supporting designers in generating creative ideas in recent decades. Databases, containing knowledge for supporting design, are often employed by the tools. Various databases are used, for instance, design repositories, ConceptNet, biological and engineering systems in structure-behaviour-function forms, and customized ones. However, some databases involve a limited amount of knowledge, some are not suitable for design, and some mainly contains prior knowledge. Besides, new knowledge emerges rapidly in nowadays fast developing world. Nowadays, to produce creative ideas for the development the innovative products and up-to-date knowledge is needed. Thereby, it is needed to explore how to employ rapidly emerged knowledge to support designers in creativity and innovation. Crowdsourcing is a model where answering open calls generate many solutions. *Goucher-Lambert and Cagan* (*2019*) have shown the use of crowdsourcing to generate inspirational stimuli to support idea generation. Social media is described as 'a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content' (*Kaplan & Haenlein, 2010*). So, such social media, as Twitter and Facebook, are considered platforms that are often used by crowds for creating knowledge. When taking advantage of crowdsourcing and social media, databases containing up-to-date knowledge created by the crowd could be constructed.

2. Problem statement

This paper aims to explore a crowdsourcing data-driven approach to construct crowd knowledge databases for innovation through supporting creative idea generation. In the approach, social media will be used as platforms to crowdsource knowledge for producing the databases.

3. Methods and data

The crowd knowledge databases strive to be employed in existing computational creativity tools for improving the tools' performances and capabilities. That benefits the generation of creative ideas and leads to innovative products. The following section investigates creativity in design. *Sections 4.1 and 4.2* explore crowdsourcing and related frameworks, respectively. *Section 4.3* proposes a crowdsourcing data-driven approach based on the explorations. *Section 4.4* discusses the challenges involved in this approach, with conclusions in *section 5*.

4. Results

4.1. Design Creativity

Creativity is considered a significant element in the design, defined as producing something judged to be creative (*Amabile, 1983*). *Han et al.* (*2019*) have indicated that novelty, surprise, and usefulness are the three core factors of creativity in design. Idea generation involves the process of creating developing and communicating ideas, where ideas are fundamental elements of thoughts in visual, concrete, and abstract forms (*Jonson, 2005*). Idea generation has been considered essential to innovation (*Sarkar & Chakrabarti, 2011; Cash & Štorga, 2015*). However, idea generation, especially generating creative ideas, is challenging in new product design and development.

Creativity tools and methods are thereby developed and used to support designers in creative idea generation during the early stages of design. There exist two categories of tools for supporting creative idea generation, non-computational and computational tools. Non-computational tools, such as TRIZ (*Altshuller, 1984*), design-by-analogy (*Goldschmidt, 2001*), and the 77 design heuristics (*Yilmaz et al., 2016*), provides designers with guidelines and instructions for producing creative ideas. Nevertheless, some of the tools rely heavily on designers' knowledge, while others are challenging to master.

In recent years, computational tools that involve the use of computational techniques for supporting idea generation have been explored. These tools could produce creative prompts and provide relevant knowledge to support designers in creative idea generation more effectively and efficiently. The Retriever (*Han et al., 2018b*) prompts designers in generating creative ideas by constructing new ontologies to support reasoning by employing real-world data. The database employed in the tool is the ConceptNet (*Speer et al., 2017*), which is a machine-understandable knowledge network. The knowledge contained is mainly practical knowledge, which has limited the Retriever in constructing profoundly new ontologies for supporting idea generation. Analogy Finder (*McCaffrey & Spector, 2017*) provides users with adaptable analogous ideas for solving technical problems by conducting searches using the US patent database. However, the tool requires the users to have substantial expertise and knowledge to adapt the ideas retrieved from the US patent database employed for solving problems. Idea Inspire 4.0 of designers (*Keshwani & Chakrabarti, 2017*) generate creative ideas for solving problems via analogical design. A searchable knowledgebase is employed in the tool containing a limited number of biological systems. An automated approach has been proposed by *Keshwani and Chakrabarti* (*2017*) for populating the database.

Creativity covers two main categories, H-creativity and P-creativity (*Boden, 2004*). H-creativity refers to historical-creativity, which indicates generating ideas that are new in history. P-creativity, also known as psychological-creativity, indicates producing new ideas to the person who produced the idea. When compared with the design creativity studies at P-creativity levels, fewer studies focus on H-creativity levels. Hence, it requires to develop design creativity at H-creativity levels, investigating how to produce new ideas to a group of people, a crowd, and, ultimately, history. From a group perspective, studies such as the ones conducted by *Paulus & Dzindolet* (*2008*) and *Nijstad & Stroebe* (*2006*) have shown that collaboration has positive effects on creativity. *Paulus et al.* (*2012*) have revealed that collaborative creativity could produce better outcomes than individual creativity. That indicates that using groups could produce ideas that are better than the ones produced by individuals. Ideas produced by a group are new to the group, which are beyond P-creativity and close to H-creativity. Employing an even more significant number of people, such as a crowd, could lead to the generation of ideas belonging to the H-creativity category.

As illustrated above, databases play a significant role in nowadays computational tools. However, the databases employed by the tools have various limitations, which have negative impacts on the tools' capabilities. Besides, the use of a crowd in supporting design creativity, especially creative idea generation, could yield superior results. A crowd could be employed to produce ideas or provide knowledge for solving design problems. The ideas produced and knowledge provided by the crowd could be constructed into a crowd knowledge database to support designers in producing creative ideas to solve the design problem. Thus, it requires a new approach to creating crowd knowledge databases for computational tools to support designers in creative idea generation needs.

4.2. Crowdsourcing for Innovation

Crowdsourcing is described as a web-based creative problem-solving model, in which "a distributed network of individuals produces solutions to an open call for proposals" (*Brabham, 2008*). In the context of design, *Forbes & Schaefer (2018)* suggest that crowdsourcing is most suited to evaluation and ideation, as shown in *Fig. 1*. Later design phases require a higher skill level and are therefore are harder to "open to the crowd." Therefore, the suitability for ideation and other early design stages is a consequence of the inverse relationship between the qualified crowd size and the level of skill for contribution. For example, in concept generation, "ideas are not scrutinized on their technical rigor or feasibility" (*Daly et al., 2012; Forbes et al., 2019*). The number of those qualified to make these contributions is higher than later design phases, and therefore the crowd available in this phase is large. However, that was founded on the assumption that a more significant number of contributions result in a more successful crowdsourcing initiative. *Panchal (2015)* discusses several "modes of failure" for crowdsourcing initiatives, including "a lack of submissions," but also the result of "numerous poor-quality submissions." Therefore, it is essential to consider that while we assume that a higher number of submissions is preferable, submissions can be detrimental to the success of the crowdsourcing initiative. Examples of initiatives that use crowdsourcing for idea generation include *Goucher-Lambert & Cagan (2019)*, who have used crowdsourcing techniques to "obtain inspirational stimuli" to support designers in ideation. "Connect and Develop" from Procter and Gamble is another example described as an "organization partnership" with "the world's most innovative minds." As part of Connect and Develop, Procter, and Gamble encourage the crowd to submit product ideas and suggestions according to a theme most relevant to their organization (*Dodgson et al., 2006*). Since using crowdsourcing for idea generation, Procter and Gamble's R & D productivity increased by 60%, and 45% of new initiatives had elements discovered externally (*Dodgson et al., 2006; Forbes et al., 2019*). A final example is the DARPA crowdsourcing initiative, which awarded one million dollars to a design team, external to the organization, to create an "innovative marine tank drive train" designed to significantly improve the efficiency of tank movement (*Ackerman, 2013*). According to the *Fig. 1*, the crowdsourcing demonstrated success in many idea generation initiatives (*Forbes et al., 2019*). Including the crowdsourcing process as an element of a data-driven approach for design creativity, whereby formalizing this process, could, therefore, prove useful to designers.

There are two types of crowdsourcing; active crowdsourcing and passive crowdsourcing. Active crowdsourcing is leveraged when the crowd actively participates in a contest or call for submissions. There are four types of active crowdsourcing initiative, crowdsourcing contests, open calls with direct rewards, open calls with direct rewards and micro-tasking. *Table 1* below gives definitions and examples of these crowdsourcing initiatives.

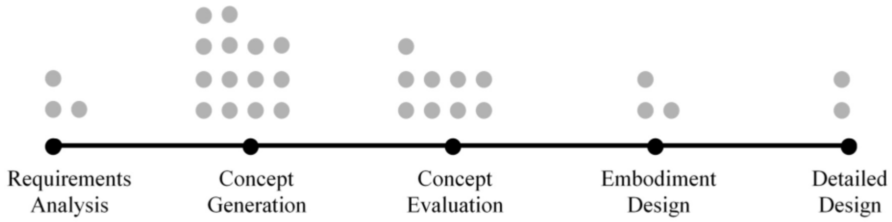

Fig. 1. Current literature's exploration of crowdsourcing in each product development phase*

**Source:* compiled based on (*Forbes & Schaefer (2018)*.

Table 1

Active crowdsourcing initiatives*

Initiative	Example	Description
Crowdsourcing contests	Gold Corp "Global Search Challenge" (*Brabham, 2008*)	Participants from different countries were encouraged to examine geologic data from Goldcorp's Red Lake Mine and submit proposals to identify potential targets with the next 6 million ounces of gold gets found. The $500,000 in prize money was offered to the 25 top finalists who identified the most gold deposits (*Brabham, 2008; Corp, 2001*)
Open calls with direct rewards	Procter & Gamble's Connect & Develop (*Dodgson et al., 2006*)	Procter & Gamble advertise their research and development needs through a crowdsourcing initiative, called Connect & Develop, in classified categories. Anyone interested in or has a solution within the advertised categories could propose their ideas by submitting through the Connect & Develop a website. The ideas get assessed by a specialized team, and reward payments can range from $10,000 to $100,000 (*Dodgson et al., 2006*)
Open calls with indirect benefits	Dell Idea Storm	With a similar setup to Connect & Develop, Dell Idea Storm is looking for ideas on their website from a community of non- experts. Contributors, however, are not rewarded financially and instead benefit indirectly from the company's implementation of the ideas in their products (*Di Gangi & Wasko, 2009*)
Micro-tasks	Amazon Mechanical Turk	Amazon Mechanical Turk is a website allowing businesses to hire participants "to perform discrete on-demand tasks that computers are currently unable to do." (*Buhrmester et al., 2011*)

*Source: compiled based on (*Panchal, 2015*).

On the other hand, passive crowdsourcing uses information from the crowd in the public domain or has been collected with permission from the crowd (*Charalabidis et al., 2014*). The used information depends entirely on the methodology applied by the data collectors, and the data content does not influence it. An example of passive crowdsourcing is Netflix's use of customer choices to supply film and TV recommendations.

Using crowd data to populate computational creativity tools is a hybrid crowdsourcing approach using both active and passive crowdsourcing. An open call with indirect rewards, an active crowdsourcing initiative, is used to encourage the crowd to share their ideas. Followed by a set method to process the data for use in a computational creativity tool, representing a passive crowdsourcing approach. Several other authors have implemented hybrid active and passive crowdsourcing approaches. For example, *Janssen et al.* (*2017*) use a hybrid approach to crowdsourcing for policymaking. They state that "combining both approaches can create the synergy. Passive crowdsourcing results can guide active crowdsourcing to avoid asking users for similar types of input".

Similarly, Charalabidis et al. (*2014*) use a hybrid approach for policymaking by "exploiting the extensive political content was continuously created in numerous Web 2.0 [technologies]". Finally, *Akshay et al.* (*2018*) use passive and active crowdsourcing for monitoring video for critical events stating that this approach "increases the feasibility of deploying continuous real-time crowdsourcing systems in real-world settings." Therefore, there is evidence of using crowdsourcing and an active-passive crowdsourcing approach for innovation in several research fields.

Despite evidence of similar successful uses of crowdsourcing, some crowdsourcing initiatives are more effective than others (*Panchal, 2015*). Ineffective crowdsourcing initiatives may invite incomplete submissions that fail to reach the required quality. A crowdsourcing initiative can also become ineffective if running the initiative exceeds the cost of an in-house team (*Brabham, 2008; Panchal, 2015*). As a result, it requires frame crowdsourcing processes. The following section presents existing crowdsourcing frameworks.

4.3. Crowdsourcing Frameworks

Crowdsourcing has emerged with the birth of the internet and with the ability to share information quickly and easily, worldwide. Social media has been a catalyst in this growth by facilitating and supporting users to create, share, and edit information, as well as build relationships through interaction and collaboration (*Mount & Martinez, 2014*). *Kemp* (*2019*) reported that there are 3.48 billion social media users in 2019, which leads to millions of posts every minute (*Forbes et al., 2019*). When an open call crowdsourcing initiative starts on social media, therefore, potential participants can be reached with quick and easy ideas submission. Preventing crowdsourcing failure, when leveraging social media, requires a methodical approach. Before presenting a new crowdsourcing social media framework for computational creativity, the authors explored existing research.

Crowdsourcing frameworks are most prevalent in the field of product design and development. *Niu et al.* (*2019*) present a framework for crowdsourcing in product development, guiding users through critical crowdsourcing decisions. *Panchal* (*2015*) also presents a framework for crowdsourcing in product development, providing a four-step approach to crowdsourcing application. This framework includes three key steps; selecting crowdsourcing initiatives, making a design decision, and incentive design. Panchal also provides further detail regarding "incentive design" by presenting a game-theoretic model for managing crowd participation. Similarly, *Abrahmason et al.* (*2013*) present an "Incentives Mix Framework" for understanding crowd participation, and *Cullina et al.* (*2016*) and *Gerth et al.* (*2012*) provide in-depth research on finding the "qualified crowd" in crowdsourcing contests. Finally, *Kittur et al.* (*2011*) consider the crowdsourcing of Human Intelligence Tasks (HITs) and "provide a systematic and dynamic way for breaking down tasks into subtasks and manage the flow and dependencies between them."

In other fields, few authors have presented a crowdsourcing framework for their domain. *To & Shahabi* (*2018*) propose a crowdsourcing framework for "protecting worker location privacy in spatial crowdsourcing," *Liu* (*2014*) present a "crisis crowdsourcing framework" for "designing strategic configurations of crowdsourcing for the emergency management domain" and *Chen et al.* (*2009*) present a "QoE evaluation framework for multimedia content." These authors represent the scarcity of crowdsourcing frameworks and demonstrate the relative youth of this research topic. By creating a crowdsourcing framework for creativity, and specifically, computation, creativity is a significant contribution in an emerging literature sector. Furthermore, existing crowdsourcing frameworks are generally at a low level of abstraction, addressing, and guiding small aspects of the crowdsourcing process instead of offering high-level support. For example, *Cullina et al.* (*2016*) discuss the need to understand crowd motivation in contests, which is a single factor contributing to the successful implementation of crowdsourcing. By presenting a high-level crowdsourcing framework for computational creativity, the authors offer more holistic guidance for crowdsourcing applications.

4.4. The Crowdsourcing Data-driven Approach

As illustrated above, crowdsourcing initiatives allow varied and numerous data points to be collected from the crowd. They are particularly useful in early design phases as the prerequisite skill level for participation in these phases. This section demonstrates how crowdsourcing could acquire knowledge from a crowd to support creative design activities in new product design and development, such as idea generation and evaluation, by partnering crowdsourcing with computational creativity tools.

A novel approach using social media to crowdsourcing design knowledge for creating crowd knowledge databases is proposed, as shown in *Fig. 2*. In step 1, an open design challenge call is posted on social media, such as Twitter and Facebook. A dedicated hashtag is involved in the open call post. The hashtag will help the crowd identify the open call on social media and be used as a target to support the later data mining process. In step 2, an effective crowdsourcing method encourages the crowd to generate ideas using descriptive text to solve the design challenge in the open call.

The ideas generated are posted back on social media containing the dedicated hashtag. Data mining is conducted in the next step to retrieve posts containing the dedicated hashtag only. That will help to discard noise data, which are irrelevant to the open call. In step 4, the retrieved data are processed using natural language processing tools to extract useful words and phrases. The extracted data are used to construct crowd knowledge databases for supporting creativity and innovation in step 5. In the last step, the crowd knowledge databases constructed will be used by exiting computational design creativity tools to enhance the capabilities of the tools in supporting idea generation. For example, the Combinator (*Han et al., 2018a*) can employ the databases to produce combinational prompts associating knowledge produced by the crowd.

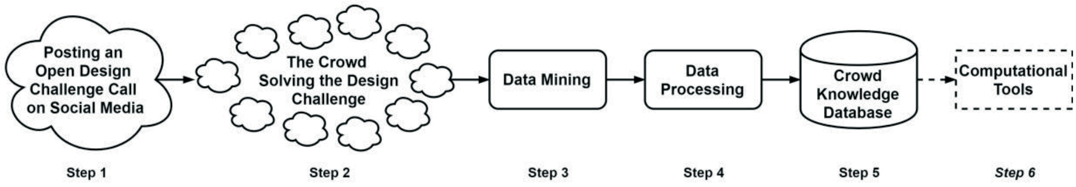

Fig. 2. The crowdsourcing data-driven approach of creating a crowd knowledge database*

*Source: compiled by Authors.

5. Discussion

After having presented the approach, this section considers the hurdles and challenges for implementation. There are three critical phases of the approach that require attention. That includes, firstly, how participation will be encouraged and managed. Secondly, how the submitted responses will be processed is significant in determining the value of ideas generated from this crowd-knowledge database. At last, it is essential to determine how the submitted responses are included as part of the computational creativity tool and whether this should differ from other databases. The third phase, regarding the use of the database, is managed by existing computational creativity tools, but the discussion (*Forbes et al., 2019*) includes the first and second phases.

When considering the management of participation, social media allows access to the most significant number of people possible, making it an effective medium for hosting passive and active crowdsourcing initiatives. The difficulty, however, is gaining active participation in these platforms. "Social media is used extensively and constantly to attract attention, and users can often be overwhelmed with online content" *(Forbes et al., 2019*). Enticing submissions, therefore, requires strategic thinking. Besides, high numbers are essential, but a wide variety is also crucial for generating innovative ideas (*Howe, 2006*). Organizations use crowdsourcing initiatives because they recognize a need to involve other perspectives beyond those of their in-house teams. Therefore, the effort must be made to increase the hashtag's exposure while limiting the "echo chamber effect" that can reduce the heterogeneity of the responses (*Colleoni et al., 2014; Forbes et al., 2019*). There is a need to manage how the hashtag is exposed to potential crowdsourcing participants to ensure text-based responses from users are useful for generating creative ideas.

Within crowdsourcing and creativity research domains, solutions to this challenge are limited. Therefore, the authors considered other research domains, such as digital marketing, to offer an understanding of how organizations can compete for social media attention while running a crowdsourcing initiative. So that to correspond with the required traits of captured data, the authors were interested in solutions to capture diverse information and solutions to capture numerous data. Regarding managing diversity, existing literature on the impact of social media on political preference offered insight. Ensuring a heterogeneous dataset meant limiting the impact of "social media bubbles" or "echo chambers" (*Zhan et al., 2016; Romero et al., 2011*), which is of significant interest in the current political climate. *Garimella et al.* (*2017*) offer a solution that could apply to crowdsource for computational creativity. They suggest when "exposing information" to users, a "symmetric difference function" could be "optimized" to limit the dominance of one piece of information in the case of two competing instances of information. In the context of ensuring diverse submissions, engaging a "symmetric difference function" could ensure that a single submission on the social platform would not influence subsequent submissions. *Dubois & Blank (2018*) also propose another solution, which suggests the ownness is on the user to limit their vulnerability to polarising online content. They demonstrate that users with "diverse interests" on social media platforms are significantly less susceptible to exposure to polarising content. Therefore, a solution to ensure heterogeneous submissions for a crowdsourcing activity could be target users with connections with a range of interests and political viewpoints.

The authors were also interested in learning how a crowdsourcing initiative could "compete for attention" on social media platforms (*Romero at al., 2011*). *Feng et al.* (*2015*) suggest garnering attention on "busy" social media platforms, information sharers need to understand how and when users become "overload with information" and respond accordingly. They show how information spread on social media can be represented by a fractional susceptible infected recovered (FSIR) model. In this case, bacteria spread is analogous to information spread and infection presents information overload (*Feng et al., 2015*). Using this model, *Feng et al.* (*2015*) suggest spreading information early in the day and early in a "social information cycle," which they describe in detail. *Iyer & Zsolt* (*2015*) suggest that to compete for attention on social media platforms, information sharers must consider the incentives users respond to for social media use in general. They then suggest embedding these incentives, such as the ability to connect with others, into the mechanism they use to spread information (*Iyer & Zsolt, 2015*). Each of these existing solutions can be considered when implementing the crowdsourcing data-driven approach. How the submitted responses will be processed is significant in determining the value of ideas generated from this crowd-knowledge database. Using texts to provoke the designers' mind in producing creative ideas has been demonstrated in several previous studies, but in various forms (*Forbes et al., 2019*). For example, *Shi et al.* (*2017*) employed network-based texts, while *Han et al. (2018a*) used combinational texts. However, the presentation form of crowd knowledge, the solutions generated by the crowd and processed by computational means in this study, still needs to be explored (*Forbes et al., 2019*).

The collection of social media data differs from data (text) used in previous studies. Crowd data may include sentimental as well as emotional aspects. That means that the process of natural language process must include a measurement of sentiment to determine the positivity and negativity of the whole text. Overall, emotionality needs to be calculated on individual text segments to indicate positive and negative text segments. Emotionality could support designers in decision-making by ensuring they have a greater understanding and further context of crowd data. For example, designers might need to avoid the design aspects of negative knowledge and enhance design features related to actual knowledge (*Forbes et al., 2019*). That might also help the computational tools in a better comprehension of the crowd knowledge database employed.

The way social media users communicate has developed beyond just text-based, which should be considered, further, to processing emotional and sentimental aspects of participant responses, "Emojis," "GIFs" and "memes" are frequently and extensively used on social media to communicate ideas. Their use means either they must be filtered and removed, or "translated" for inclusion in a crowd database. One approach to this, as shown in *Fig. 2*, includes using keywords to identify the critical idea communicated in participant responses. However, it could be the case when the critical idea is communicated in a text-based caption with an image accompaniment to bolster, as opposed to conveying, the idea. How this different use of video and image-based content is managed should be taken into consideration.

Twitter and other social media platforms are purposefully designed to encourage collaboration and interaction between users. That results in functionality elaborating and "commenting" on other responses that are considered integral to the design of these online platforms. As a result, however, the processing of participant involvement needs to recognize not only individual responses, including the hashtag but "clusters" or responses that all represent one idea (*Forbes et al., 2019*). For example, one participant may include the "crowdsourcing hashtag" to present an idea that initiates an online conversation, with further responses elaborating on or supporting the initial idea. Some of these comments may be new ideas, but others could be minor alterations or additions to the original submission. That means that including every response involved in the conversation and weighting them could equally disrupt the value of crowd data. An understanding of how collaboration occurs on social media is fundamental to procuring valuable results for idea generation (*Forbes et al., 2019*).

Utilizing crowd knowledge from social media shows great potential for supporting creativity and innovation. However, there are several research challenges, such as participation management and data processing, to overcome. Furthermore, the way social media users communicate has and will change to incorporate more media-based content. Further research is needed to solve these research challenges and recognize new opportunities in the applications of this crowdsourcing data-driven approach. The critical next research step is to conduct a case study using crowd knowledge from a specific social media platform to solve a design challenge. The authors hope to provide more insights into this new and novel data-driven computer-aided innovation approach.

6. Conclusions

Generating ideas, especially creative ones, is significant to innovation. However, it is challenging to produce creative ideas. Many computational support tools are developed to assist this process, but available databases constrain the current solutions. Lacking knowledge in terms of quantity and variety is one of the main issues of the databases. Besides, knowledge collection has been considered a time-consuming and frustrating activity (*Darma et al., 2020*). Crowdsourcing is a model for creative problem-solving which uses the knowledge produced by a distributed network of individuals, also known as a crowd. Social media, which allows creating and exchanging content created by users, is often employed to generate and share knowledge.

Thus, the authors of this paper have proposed a novel data-driven approach utilizing social media to crowdsourcing knowledge to construct databases for computational tools to support creative idea generation, ultimately leading to innovation. The databases constructed are called the crowd knowledge databases populated by providing and distributing open design challenge calls with responses using unique hashtags for identification. Data mining and natural language processing are used in the construction process to retrieve and extract data, respectively. The crowd knowledge databases can then be implemented into existing and future computational tools to enhance their performances. Using the Combinator (*Han et al., 2018a*) as an example, the tool could associate crowd knowledge from the database to produce new combinational prompts, which are new to the crowd, for stimulating users' creative minds. The data-driven approach proposed has implied its value of utilizing some of the most used and data-rich platforms available to achieve innovation.

However, some challenges need to be solved to realize the crowdsourcing data-driven approach. This paper discusses how to manage participation in social media and how to process various types of information. Several participation management methods, such as information spreading and incentives, as well as several information processing issues, such as sentiment measurements and collaboration, understands, are indicated. Further research is required to explore these challenges and to overcome them, to fully employ the proposed crowdsourcing data-driven approach in computational support tools for innovation. This paper has thereby shown a new research direction in using crowdsourcing data to support innovation, contributing to the computer-aided innovation research area. The authors have planned to conduct a case study of solving a design challenge using the crowd knowledge from a specific social media platform, such as Twitter, in their next study to provide more valuable insights.

7. Funding

This study received no specific financial support.

8. The competing interests

The authors declare that they have no competing interests.

References

Abrahamson, S., Ryder, P., & Unterbeg, B. (2013). *Crowdstorm: the future of innovation, ideas, and problem solving*. New Jersey: John Wiley & Sons.

Ackerman, S. (2013). This Is the Million-Dollar Design for Darpa's Crowdsourced Swimming Tank. Retrieved July 17, 2020 from https://www. wired. com/2013/04/darpafang-winner

Akshay, R. K., & Walter, S. L. (2018). Plexiglass: Multiplexing passive and active tasks for more efficient crowdsourcing. *Sixth AAAI Conference on Human Computation and Crowdsourcing*.

Altshuller, G. S. (1984). *Creativity as an exact science: The theory of the solution of inventive problems*. Amsterdam: Gordon and Breach Publishers.

Amabile, T. M. (1983). *The social psychology of creativity*. New York: Springer-Verlag.

Bertola, P., & Teixeira, J. C. (2003). Design as a knowledge agent: How design as a knowledge process is embedded into organizations to foster innovation. *Design Studies*, *24*(2), 181-194.

Boden, M. A. (2004). *The creative mind: Myths and mechanisms (Vol. 2)*. London: Routledge.

Brabham, D. C. (2008). Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. *Convergence*, *14*(1), 75-90.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data?". *Perspectives on psychological science*, 6(1), 3–5.

Cash, P., & Štorga, M. (2015). Multifaceted assessment of ideation: using networks to link ideation and design activity. *Journal of Engineering Design*, 26(10-12), 391-415.

Charalabidis, Y. N., Loukis, E., Androutsopoulou, A., Karkaletsis, V., & Triantafillou, A. (2014). Passive crowdsourcing in government using social media. *Transforming Government: People, Process and Policy*, *8*(2), 283-308.

Chen, K. T., Wu, C. C., Chang, Y. C., & Lei, C. L. (2009). A crowdsourceable QoE evaluation framework for multimedia content. *Proceedings of the 17th ACM international conference on Multimedia* (pp. 491-500). ACM.

Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. *Journal of Communication*, 64(2), 317-332.

Corp, G. (2001). The goldcorp challenge winners. Retrieved July 11, 2020 from http://www.goldcorpchallenge.com/challenge1/winnerslist/chaleng2.pdf.

Cullina, E., Conboy, K., & Morgan, L. (2016). Choosing the right crowd: An iterative process for crowd specification in crowdsourcing initiatives. *System Sciences (HICSS) 49th Hawaii International Conference* (pp. 4355-4364). IEEE.

Daly, S. R., Yilmaz, S., Christian, J. L., Seifert, C. M., & Gonzalez, R. (2012). Design heuristics in engineering concept generation. *Journal of Engineering Education*, *101*(4), 601-629.

Darma, D. C., Ilmi, Z., Darma, S., & Syaharuddin, Y. (2020). COVID-19 and its Impact on Education: Challenges from Industry 4.0. *Aquademia, 4*(1), ep20025.

Di Gangi, P. M., & Wasko, M. (2009). Steal my idea! organizational adoption of user innovations from a user innovation community: A case study of dell ideastorm. *Decision Support Systems*, *48*(1), 303–312.

Dodgson, M., Gann, D., & Salter, A. (2006). The role of technology in the shift towards open innovation: the case of Procter & Gamble. *R&D Management*, *36*(3), 333-346.

Dubois, E., & Blank, G. (2018). The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, Communication & Society*, *21*(5), 729-745.

Feng, L., Yanqing, H., Baowen, L. H., Eugene, S., Shlomo, H., & Braunstein, L. A. (2015). Competing for Attention in Social Media under Information Overload Conditions. *PloS, 10*(7), e0126090.

Forbes, H. L., & Schaefer, D. (2018). Crowdsourcing in product development: Current state and future research directions. *Proceedings of the DESIGN 2018 15th International Design Conference*. Dubrovnik, Croatia: The Design Society.

Forbes, H., Han, J., & Schaefer, D. (2019). Exploring a Social Media Crowdsourcing Data-Driven Approach for Innovation. Proceedings of the International Conference on Systematic Innovation. *The 2019 International Conference on Systematic Innovation*. Liverpool, UK.

Garimella, K., Gionis, A., Parotsidis, N., & Tatti, N. (2017). Balancing information exposure in social networks. *31st Conference on Neural Information Processing Systems (NIPS 2017)*. Long Beach, CA, USA.

Gerth, R. J., Burnap, A., & Papalambros, P. (2012). *Crowdsourcing: A primer and its implications for systems engineering*. Michigan University Ann Arbor.

Goldschmidt, G. (2001). Visual Analogy - a Strategy for Design Reasoning and Learning. In G. Goldschmidt, *Design Knowing and Learning: Cognition in Design Education* (pp. 199-219). Oxford: Elsevier Science.

Goucher-Lambert, K., & Cagan, J. (2019). Crowdsourcing inspiration: Using crowd generated inspirational stimuli to support designer ideation. *Design Studies*, *61*(1), 1-29.

Han, J., Forbes, H., & Schaefer, D. (2019). An Exploration of the Relations between Functionality, Aesthetics and Creativity in Design. *Proceedings of the Design Society: International Conference on Engineering Design*, *1*(1), 259-268.

Han, J., Shi, F., Chen, L., & Childs, P. R. (2018). A computational tool for creative idea generation based on analogical reasoning and ontology. *Artificial Intelligence for Engineering Design*, *32*(4), 462-477.

Han, J., Shi, F., Chen, L., & Childs, P. R. (2018). The Combinator – a computer-based tool for creative idea generation based on a simulation approach. *Design Science*, *4*, e11.

Howard, T. J., Culley, S., & Dekoninck, E. A. (2011). Reuse of ideas and concepts for creative stimuli in engineering design. *Journal of Engineering Design*, *22*(8), 565-581.

Howe, J. (2006). The rise of crowdsourcing. *Wired Magazine*, *14*(6), 1-4.

Iyer, G., & Zsolt, K. (2015). Competing for attention in social communication markets. *Management Science*, *62*(8), 2304-2320.

Janssen, M., Konopnicki, D., Snowdon, J. L., & Ojo, A. (2017). Driving public sector innovation using big and open linked data (BOLD). *Information systems frontiers*, *19*(2), 189-195.

Jonson, B. (2005). Design ideation: the conceptual sketch in the digital age. Design Studies, 26(6), 613-624.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*(1), 59-68.

Kemp, S, Hootsuite. (2019). Digital Report 2019: Essential Insights into How People Around the World Use the Internet Mobile Devices. Retrieved July 8, 2020 from https://www.juancmejia.com/wp-content/uploads/2019/03/Digital-2019-WeAreSocial-y-HootSuite.pdf

Keshwani, S., & Chakrabarti, A. (2017). Towards automatic classification of description of analogies into SAPPhIRE constructs. *Research into Design for Communities, 2*, 643-655.

Kittur, A., Smus, B., & Khamkar, S. K. (2011). Crowdforge: Crowdsourcing complex work. *Proceedings of the 24th annual ACM symposium on User interface software and technology* (pp. 43-52). ACM.

Liu, S. B. (2014). Crisis crowdsourcing framework: Designing strategic configurations of crowdsourcing for the emergency management domain. *Computer Supported Cooperative Work*, *23*(4-6), 389-443.

McCaffrey, T., & Spector, L. (2017). An approach to human–machine collaboration in innovation. *Artificial Intelligence for Engineering Design*, *32*(1), 1-15.

Mount, M., & Martinez, M. G. (2014). Social Media: A Tool for Open Innovation. *California Management Review*, 56(4), 124-143.

Nijstad, B. A., & Stroebe, W. (2006). How the Group Affects the Mind: A Cognitive Model of Idea Generation in Groups. *Personality and Social Psychology Review*, *10*(3), 186-213.

Niu, X. J., Qin, S. F., Vines, J., Wong, R., & Lu, H. (2019). Key crowdsourcing technologies for product design and development. *International Journal of Automation and Computing*, *16*(1), 1-15.

Panchal, J. H. (2015). Using crowds in engineering design–Towards a holistic framework. *Proceedings of the 20th International Conference on Engineering Design (ICED 15), 8: Innovation and Creativity*, pp. 041-050. Milan, Italy.

Paulus, P. B., & Dzindolet, M. (2008). Social influence, creativity and innovation. *Social Influence*, *3*(4), 228-247.

Paulus, P. B., Dzindolet, M., & Kohn, N. W. (2012). Chapter 14 - Collaborative Creativity: Group Creativity and Team Innovation. In *Handbook of Organizational Creativity*. San Diego: Academic Press.

Romero, Mauricio, D., Meeder, B., Barash, V., & Kleinberg, J. (2011). Maintaining ties on social media sites: The competing effects of balance, exchange, and betweenness. *Proceedings of the Fifth International Conference on Weblogs and Social Media*. Barcelona, Spain.

Zainurossalamia, S., Darma, D. C., Kasuma, J., Ratnasari, S. L., & Tasențe, T. (2020). Apparatus performance as mediation of creativity and innovation towards the successful application of e-kelurahan. *European Journal of Human Resource Management Studies, 4*(2), 108-126.

Sarkar, P., & Chakrabarti, A. (2011). Assessing design creativity. *Design Studies*, *32*(4), 348-383.

Shi, F., Chen, L., Han, J., & Childs, P. (2017). A Data-Driven Text Mining and Semantic Network Analysis for Design Information Retrieval. *Journal of Mechanical Design*, *139*(11), 111402-111414.

Siddharth, L., & Chakrabarti, A. (2018). Evaluating the impact of Idea-Inspire 4.0 on analogical transfer of concepts. *Artificial Intelligence for Engineering Design*, *32*(4), 431-448.

Speer, R., Chin, J., & Havasi, C. (2017). ConceptNet 5.5: an open multilingual graph of general knowledge. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*. San Francisco, California.

To, H., & Shahabi, C. (2018). Location privacy in spatial crowdsourcing. *In Handbook of Mobile Data Privacy* (pp. 167-194). Springer, Cham.

Ullman, D. G. (2010). *The mechanical design process: Part 1*. New York: McGraw-Hill.

Yilmaz, S., Daly, S. R., Seifert, C. M., & Gonzalez, R. (2016). Evidence-based design heuristics for idea generation. *Design Studies*, *46*(C), 95-124.

Zhan, L., Yongqiang Sun, N. W., & Xi, Z. (2016). Understanding the influence of social media on people's life satisfaction through two competing explanatory mechanisms. *Journal of Information Management*, *68*(3), 347-361.