Text Mining Approach for Identifying Product Ideas and Trends Based on Crowdfunding Projects

David Boye – De Montford University

Sercan Ozcan & Oluwatobi Fajana – University of Portsmouth

ABSTRACT

Current research and development investments and innovation developments show that there are more failed innovation attempts than ever and there are numerous occasions where companies need to be highly selective by assessing customers' needs or preferences. Crowdfunding is a great approach for this purpose, where investments are made if there are funders who are interested in the proposed innovations. As crowdfunding platforms have proved to assist in the introduction of innovations, the details of these projects can be a great inspiration for others. The aim of this study is to integrate crowdfunding projects into innovation and product developments as a source of potential ideas and trends. We use a text mining approach to analyse 8,021 crowdfunding projects from the period 2009–2018. In this study, we cluster these projects into nine innovation areas desired by consumers. Through our methodological approach, we examine the linkage between associated text, features of projects, and funding of projects to uncover emerging product features that illustrate the desires of consumers. Our proposed model offers theoretical contributions to the innovation processes, especially the fuzzy front end with an open innovation approach. We provide practical contributions by revealing crowdfunding platforms as a means to gain insights into product elements for the design of new products and services.

Keywords: crowdfunding, customer-centric innovation, text mining, fuzzy front end, innovation process

I. Introduction

Success in innovation management requires a process and quality ideas; however, the failure rate of newly introduced innovations is relatively high [1]. The fuzzy front end (FFE) phase is considered to have an enormous effect on the outcomes of the innovation process. This phase is known to be unstructured, as ideas can emerge and evolve in an informal, chaotic and creative intensive process. Although the FFE phase is the first stage of the innovation process, initiation of the phase can be based either on problems induced by market pressures or on identifying opportunities [2, 3].

Organizations specializing in identifying new technological opportunities tend to experience higher growth than those guided by market-induced pressures [4]. Thus, opportunity

identification remains crucial to preserve the sustainable growth of organizations. Organizations can either focus on incremental modifications to existing product lines or those radically new to world products. However, it is seen that organizations have largely focused more on incremental product modifications due to the need for support and lack of experienced New Production Development (NPD) teams and tools when dealing with uncertainties concerning discontinuous technologies [5]. Hence, identifying a systematic approach to start the phase is of importance [6].

Due to this importance, organizations utilize traditional customer-centric analysis such as focus groups, product trials, surveys and so on. Still, these methods are proving less viable and static in fast-paced environments. This is due to the level of bias, their costly nature, and the lack of customers possessing adequate information [7]. According to Kutvonen and Torkkeli [8], adopting an open innovation viewpoint in the fuzzy front end can enable organizations to be mindful of opportunities in the external landscape, thereby reducing unprofitable kill decisions and concept development waste. However, the authors highlight the lack of research in external concept exploitation. Researchers have proposed innovative methods for discovering opportunities using various techniques and data sources. Several studies have suggested using external data sources such as online reviews, social media, patents, and publications [9, 10, 11, 12].

Kim and Bae [9] proposed a novel approach to forecasting promising technology using patent analysis. Similar to the present study. Zhang, Rao and Feng [10] examined consumer online reviews to show the correlation between online reviews and phone improvement. Wang and Chen [11] proposed a patent mining approach to identify technological opportunities, while Ozcan et al. [12], based on Twitter data, proposed an end-to-end framework that would reveal clusters of ideas with a word network map. However, a significant weakness in the literature is the reliance on data sources that are either too industry oriented, which causes over-dependence

on forecasting methods (i.e. patents), or too customer-centric, which lacks the involvement of industrial experts (social media) [13]. Furthermore, few studies propose how this analysis can aid the FFE phase. The front-end literature acknowledges the importance of a systematic approach, with studies examining how to create radical ideas successfully [14]; the relevance of divergent thinking and market visioning [5]; user experience-driven idea generation [6]; the importance of strategies for selecting early stage opportunities [15]; striking the right balance in managing the quantity and novelty of ideas [16]; the influence of organizational structures and cultures [17]; the clustering of front end innovation tools based on macro and micro parameters [18]; and idea generation through morphology analysis [19].

Overall, the front-end literature still lacks insight into understanding customer preferences, which results in the development of solutions that do not meet customer needs. According to Achiche et al. [18], uncertainties concerning the market and environment are inevitable. Furthermore, Zhang and Doll [20] explained that during the FFE phase of innovation developments, uncertainties arise from lack of clarity considering what customers value. Hence, a novel ideation approach for the FFE phase is needed to identify valuable and customer-centric innovation ideas. Accordingly, this study proposes that external concept exploitation can be achieved to identify product opportunities based on the direct relationship between acquired product funding and the products themselves using crowdfunding data in customer-centric product analysis approaches to minimize or eliminate the problems mentioned above.

Crowdsourcing platforms have become sources of ideas and product design. Crowdfunding, a subset of crowdsourcing [21], enables consumers to participate in the innovation process by contributing funds and feedback to upcoming ideas and innovation [22]. According to Butticè and Ughetto [22], crowdfunding research ranges between three main categories: actors, crowdfunding characteristics, and campaigns. Crowdfunding campaigns can also become a

medium beyond financing and more of an informational mechanism [23]. Innovations such as GoPro, Pebble, and so on, which consumers funded, reveal inbuilt features that signal the desires of consumers and are beneficial to organizations during the ideation process to achieve innovations with superior performance in markets [24]. Sourcing of user-generated content and perceptions of technology developments can highlight precise trends in emerging technologies [25]. This study aims to develop an opportunity identification framework utilizing crowdfunding platforms to enable organizations to gain competitive intelligence and ideation for product development. Crowdfunding data from 2009 to 2018 containing innovative projects, project descriptions, funding amounts, project year and title were compiled. The primary approach was to create a text mining co-occurrence and clustering framework that can aid opportunity identification based on consumer funding of projects.

This study has theoretical, practical and methodological contributions. From a theoretical perspective, we contribute to the open innovation and innovation process models and theories by integrating the pre-phase, phase zero, and phase one steps with an outside-in approach where innovation ideas and market trends can be captured from crowdfunded project data. We add to the studies of Spitsberg et al. [26], Christensen et al. [27], Lee and Sohn [28], Khademi et al. [29] and Ozcan et al. [12], where innovation ideas are retrieved from different sources such as social media and patent and crowdfunding data. Our study contributes to these approaches by introducing a full-scale end-to-end innovation development process where the innovation ideas are retrieved from those funding the crowdfunded project. Hence, our approach minimizes the investment risk by focusing on the opportunities that are invested by end users. This study also has a methodological contribution by providing an optimized data retrieval approach for sourcing ideas from crowdfunding data. The study makes a practical contribution by illustrating the funded technologies with examples such as 3D printing technologies. The structure of the paper is as follows. Firstly, we provide a literature review on product innovation and

crowdfunding, text mining techniques and technology trends. Secondly, we present the crowdfunding mining process. Thirdly, we provide crowdfunding ideation results. Finally, we conclude with key findings, implications and suggestions for future research.

II. Literature Review

Innovation is essential for the fastest growing firms. Many new products do not necessarily displace existing products but rather create new markets by solving numerous unsolved consumer problems. The basics of innovation and new ideas lie in published innovations and organizations utilizing appropriate procedures to achieve successful sustainable outcomes [12, 25, 30]. This section examines the literature on the fuzzy front-end phase, ideation and new product development process. Then we examine the literature on data sources and text mining techniques to support organizational activities.

A. Fuzzy Front End and Opportunity Ideation

Developing new product concepts and technologies is exclusively the result of organizations' research and development (R&D) and undergoing new product development (NPD) processes. A new technology is said to be the combination of existing technologies in a problem area, causing a paradigm shift in markets [31]. To assist with uncovering new technologies, traditional methods such as category appraisals, interviews, questionnaires, customer reviews, focus groups, user experiments, user toolkits, and lead user techniques are used to collect and provide insights during the ideation and design phases of innovation processes [7, 13]. Although R&D is important, research shows that established organizations can have limited vision compared to start-ups. This can be due to a reduction in R&D spending and continuous re-use of the same pathways, thereby neglecting valuable knowledge and the possibility of success in new markets [32]. Generally, the NPD process consists of phases in which an idea

is transformed into an outcome for commercial purposes by involving internal and external stakeholders [33], with many studies uncovering underlying stages and phases [34, 35, 36]. Table I illustrates several NPD processes uncovered with their underlying stages and phases.

Table I: Breakdown of NPD Process.

	Phases					
Authors	1	2	3	4	5	6
Koen et al. [34]	Fuzzy front end: Idea generation and assessment	Fuzzy front end: Concept development, product planning	Design: Development	Design: Prototypes and pilot tests	Commercialization: Production, market introduction and penetration	
Cooper [35]	Discovery stage	Scoping	Build a business case	Development	Testing and validation	Launch
Chesbrough [36]	Fuzzy front end Input, In-sourced ideas and technology	Development of I/P source for development and I/P licensing	Commercialisation Products in sourced for scale-up Technology spinoffs			

The first phase of the NPD process is the FFE phase. Studies have proposed that this phase is less structured and concerns the generation and selection of quality ideas that are diffused into subsequent stages [37]. Organizations must carefully select valuable ideas that reduce failure at the end of the NPD process [38]. Liu, Goulding, and Brailsford's [39] study describes an idea as combining old elements into a new one and applying this to new unforeseen problems. An idea is a concept for a course of action, and in the case of innovation, ideas are a source of development for beginning the FFE process [31]. However, organizations can also focus on identifying and shaping opportunities based on their capabilities [40].

Innovations can be initiated based on either an opportunity concerning technological advancement or problems that are market- or customer-induced [2]. According to Wyrtki, Röglinger and Rosemann [13], organizations that focus on detecting opportunities experience higher growth than those guided by problems. These opportunities are created by competitive defects, such as unequal information dissemination or new knowledge [41]. This necessitates enterprises constantly monitoring such flaws to detect and capitalize on possibilities. As a

result, firms must monitor internal and external trends to see opportunities to generate new creative ideas [42]. Studies have examined varying opportunity typologies [43], along with opportunity identification, individual entrepreneurs [44] and sources of opportunities [45]. Moon, Han and Kwahk [46] proposed the use of scenarios and user information during opportunity identification phases to uncover creative ideas for radical innovation. Koen, Bertels and Kleinschmidt's [47] survey of US-based companies uncovered that opportunity identification, idea enrichment and concept definition are essential for success concerning incremental innovation, while understanding both disruptive and existing markets and leveraging emerging technologies were pivotal for the success of radical innovation. Loanzon et al. [7] examined netnography as a fuzzy front-end tool and uncovered the limiting nature of focus groups and online surveys due to the shortening of time concerning customer responses, which can present a bias [18]. Overall, studies acknowledge that an opportunity-led FFE phase offers better potential by yielding incremental and radical innovation. However, in the literature, there is less emphasis on methods and tools to support opportunity identification and analysis to enable the organization to sense and seize opportunities [3]. In this study, we adopt the notion that opportunities exist in a business or technology gap between a current situation and an envisioned future, whereby opportunities reside in previously undetected product complaints, wishes, expectations, and requirements.

Several researchers state that ideas developed from a deep understanding of customer needs through internal or external sources may also increase the quality of produced ideas [36]. Although techniques such as a web-enabled idea bank and lead user techniques are beneficial, they are not always successful [6, 48]. Hence the need to complement traditional ideation approaches with digital sources such as corporate online ideation platforms and crowdsourcing to open an array of possibilities for organizations to establish connections with consumers' real-life needs [49]. According to Dahlander and Piezunka [49], either utilizing proactive attention

during internal development and presenting of concepts to consumers to encourage debates or utilizing reactive attention such as paying attention to suggestions from consumers to signal they are being listened to, can support bottom-up collaborations. A deeper look into research on crowdsourcing reveals crowd involvement during crowdsourcing modes (collaboration or competition) to source ideas and funding for innovation activities [50]. Crowdfunding has become an increasingly novel means of generating finance, offering a risk-free way for entrepreneurs to create product awareness to promote innovative ideas and mechanisms to disclose consumer opinion on an idea or product before launch [51]. Crowdfunding platforms can entail many novel projects that are funded by backers driven by emotional and social criteria rather than just financial gain [52]. A significant feature that attracts these backers to projects is the product description section, which contains product features and a project summary. This tends to lead to the likelihood of attaining funding [51]. According to Lee and Sohn [28], project campaigns consist of crowd comments and product features that can be sourced and utilized by organizations. Companies can examine a product's core to understand its features concerning the success and diffusion of any new product or technology. Crowdfunding platforms can also serve as an informational mechanism for product opportunity identification due to consumers' contributions towards specific prototypes or products [53]. This study applies a text mining approach to illustrate how crowdfunding platforms can be a data source for identifying trending technology components of consumer-centred innovations.

As innovation can be complex, a means to assist the sourcing and integration of ideas to satisfy consumer desires is necessary to provide solutions during innovation processes. Eling and Herstatt [3] proposed that there are significant gaps in successful FFE management. Studies have examined the FFE phase by taking a process perspective, identifying the pillars and proposing methods for FFE management. Compared to problem-centred pressure literature, research on opportunity identification and organizations seizing opportunities is quite lacking.

In relation to open innovation, it is necessary to complement organizational boundaries and internal capabilities with external collaboration with stakeholders to boost innovation competencies [8]. Examining the literature on NPD and crowdsourcing, most studies focus on the benefits and application of crowdsourcing for innovation activities, but with less focus on the FFE phase of the innovation process. Furthermore, it is identified that traditional methods, such as surveys, focus groups and so on, have a negative outlook due to the static nature of the information collected, their costly nature, and the lack of consumers possessing adequate information [7]. Although there are studies that contribute to the literature by proposing the importance of striking the right balance between managing the quantity and novelty of ideas, user experience-driven idea generation and the importance of market envisioning [6, 15, 16], there is need for more insight into understanding customers' preferences to develop solutions that meet their needs. As a result, we suggest a platform-based mining model for FFE opportunity identification and ideation centred on the crowdfunding environment to enable the sourcing and clustering of large-scale data to support the concerns above.

B. Text Mining, Opportunity Identification and Ideation

The proliferation of Web 2.0 technologies have made potential sources of innovative ideas ubiquitous. Knowledge can be uncovered by applying statistical analysis techniques to databases to discover patterns and technical information automatically. One of the many techniques is text mining. This is a prevalent knowledge discovery technique for analysing unstructured or semi-structured data by putting a set of labels on documents and characterizing documents according to keywords extracted through algorithms [54]. Existing text mining techniques range from text segmentation to topic identification, term association, information mapping, and visualization cluster generation. In many literature domains, the visualization clustering technique makes it easier to identify the hidden structure of information and reduces

the time consumed reading to identify trends and phases of technologies [31]. Text is a common medium for exchanging information, and studies of publications, computer logs, emails, speeches, and patents can reveal relationships between technological advances, economic progress, and innovation developments within national and global contexts [12]. Table II shows existing studies utilising various data sources to identify trends and uncover trending topics and emerging technologies over the years.

Table II: Breakdown of Data Sources.

Data Source	Authors	Purpose
Scientific Publications	[13, 29]	• To understand and review research output to identify and uncover novel technologies • To identify hot topics and illuminate themes in industries • To examine and predict future trends within fields of science
Patent Documents	[11, 55]	• To monitor the relatedness in the pairs of keywords to monitor emerging technologies • To examine technological linkage using a proposed linkage model • To create technology maps and forecast emerging technology areas using growth curves and scenario planning
Social Media	[12, 27]	• To detect and identify the commercial potential of emerging technologies • To collect data to utilized in the initial stages of product design lifecycle • To uncover business value and reveal attitudes towards brands • To identify factors for market prediction

The various data sources [12, 29, 55] have their own strengths and weaknesses due to biases towards certain business models, products or services [7]. Patent documents can provide evidence as to which products enable market success but can encourage competitor imitation from an industrial perspective [55]. Scientific publications can enable the development of new product features that can influence customer behaviour, but these inventions are not largely utility driven and are based on little or no market knowledge, thereby requiring customer education [56]. Social media are beneficial for identifying information on customers' concerns and products, but the data can be of low quality due to a less expert community [12]. Overall, an increase in the number or quality of data sources can support an organization to realize its technical strength [57]. Previous studies have revealed that the application of text mining techniques can be of assistance during decision-making. Problem areas such as identifying technological features have implemented such techniques and can be grouped into categories such as expert qualitative, indicators analysis, and citation analysis. These approaches explore

technologies' newness and growth and identify core product features [54]. Specifically, these approaches can be used for the FFE phase, specifically for ideation and opportunity identification. However, Guo et al. [57] argued that these enabling techniques do not always help companies arrive at new ideas; the authors proposed that companies can further realize their technical strength and competitive situation by using data sources such as patent documents. Although 1 per cent of patents are genuine originations, ideas and conceptions with novelty methods can support the process of product innovation. Hence, effective ideation mechanisms are required to retrieve ideas from a variety of sources. Also, the reliability of these approaches needs to be increased. Some studies show how helpful knowledge extraction on user behaviour, opinion and trends can be done using text mining techniques. Regarding product ideation, the literature consists of studies that identify features of products from either the same domain type or across domains. Product designers can use these for purposes such as categorizing customer requirements and ranking existing features based on importance from a customer perspective [58]. Monitoring core technological components and product features can serve as a basis for further research, such as identifying technology opportunities, forecasting, and extracting technological intelligence [59].

According to open innovation-based data analysis studies, the difference in search strategies can influence organizations' ability to achieve a different level of novelty in innovation activities [60]. Lee et al. [33] utilized the term frequency-inverse document frequency (TD-IDF) technique and sentiment analysis using term and non-term features on a platform (MyStarbucksIdea.com) to develop classification models to calculate the adoption probability of each idea in an open innovation environment. This study proposes crowdfunding platforms as an emerging data source to assist organizations in monitoring technology development and market trends. Christensen et al. [27] proposed a machine learning technique to identify user-

generated content containing an idea from the Lego online community by retrieving and manually rating labelled texts.

Looking deeper into opportunity identification, Seo et al. [61] proposed a systematic patent approach to identify potential product opportunities by reflecting the target firm's internal capabilities and analysing the firm's existing products as a basis for developing new products. Wang and Chen [11] proposed a patent mining approach to identify technological opportunities. For this purpose, they utilized natural language processing and latent semantic analysis to determine outlier patents for potential opportunities. Kim and Lee [55] examined technology service-based opportunity identification using quantitative analysis of business method patent documents. The authors developed a portfolio map-based co-classification and interrelationship matrix uncovering potential, attractive, separated and saturated opportunities.

Studies such as those by Christensen et al. [27] and Ozcan et al. [12] utilize social media to support the FFE phase by retrieving innovation ideas, but these sources may not be as reliable as product- or project-specific data. Although the generic community can be valuable, it is argued that consumers often may not know what they want [62]. However, there is a necessary role for technology in opportunity identification and the use of data analytics to support the phase [63]. In addition, there is a need for more research examining hardware projects [28]. Overall, we identify that the majority of data sources used by scholars are from an industrial perspective focusing on patent and publication data. However, these studies do not take into account analysis in a more customer-centric environment. Moreover, previous methods have addressed evaluating consumer perception-based tweets or organizations' patent database without considering customers' willingness to invest in potential technologies, which can indicate customer preferences for technologies that can satisfy their needs. Hence, a customer-centric and more reliable approach is needed for the FFE phase of innovations. Regarding

crowdfunding platforms, projects funded by consumers present more reliable data as there is already an approved trend based on the crowd's donations to and purchases of proven designs and products. Considering the limitations in the literature, a crowdfunding-based opportunity identification approach would add value to this field, and hence a new approach needs to be introduced.

C. Conceptual Framework

According to the aim of this study, crowdfunding campaigns could be an excellent source for organizations to identify ideas for the next generation of innovations. The analysis of funded crowdfunding projects would have an advantage over other data sources, leading to better consumer acceptance and commercialization of the invention. According to Kutvonen and Torkkeli [8], organizations can benefit from external technology exploitation through either external knowledge acquisition or external concept acquisition. This exemplifies the synthesis between open innovation and the FFE phase. Hence, in this study, we will analyse these data, supporting the FFE (opportunity identification) phase. The examined projects are analysed according to customers' needs and desires and are grouped according to the type of innovation, which can range from market innovation to incremental innovation, technical innovation and radical innovation [62]. According to Herstatt and Verworn [64], the fuzzy front end can be divided into three phases: pre-phase zero, phase zero, and phase one. Pre-phase zero relates to the opportunity to generate ideas in the technical area that should align with existing projects and the project portfolio. Phase zero involves identifying preliminary market segments, customer needs, and business prospects that help define a product. Phase one relates to the business, technical feasibility, and project planning assessment to achieve a clear product concept and execution of the new product. It is proposed that management of the early phases should be adopted with a wide range of uncertainties and the important need for flexibility [64]. Start-ups can interpret signals between consumers and their relationship with products to

identify opportunities by utilizing knowledge gained from interacting with customers. This process strongly affects a firm's direction and often significantly improves the new venture's ability to develop new and distinctive products [51]. Figure 1 illustrates the conceptual framework of this process.

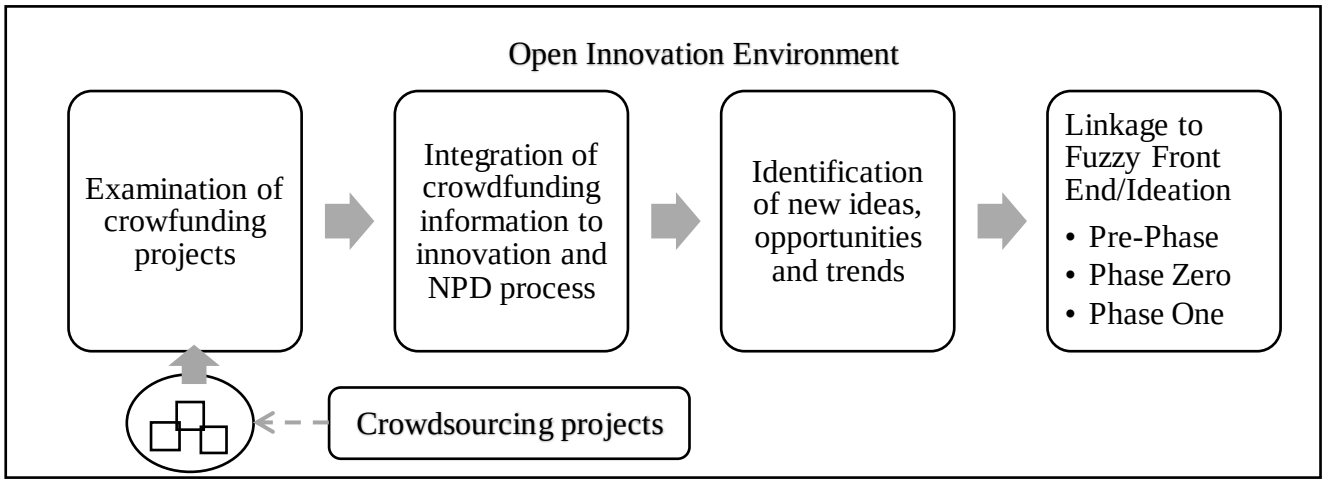

Fig. 1. Conceptual framework based on Kutvonen and Torkkeli [8] and Herstatt and Verworn [64]

III. Research Methodology

This section describes the dataset, data collection method, visualization, and interpretation for the proposed platform-based mining framework.

A. Dataset

The dataset consisted of 21,343 projects from 2009 to 2018, collected using a data retrieval approach from the website pages of a crowdfunding platform (Kickstarter). The search query retrieved page elements ranging from project title to project description, project links, project target goal, project amount reached, project category and project year. The retrieved projects were all semi-structured and derived from categories such as health, communication, energy,

green technology, and so on. We followed an ethical retrieval approach (robots.txt protocol), a set of principles for how web crawling should operate. Our process led to no service disruption or personal data use [65].

B. Research Method

The research methodology to arrive at the term-based maps and idea identification is illustrated in Figure 2. This section provides the research method's data retrieval, analysis, reliability and robustness. This study uses the terms "frequency-inverse document frequency" and "noun phrasing method" to identify technological components. Crowdfunding platforms represent a rich data source for technology categories. To simplify the process, text mining enabled us to identify the clusters of technologies within the dataset. The platform categories supported the clustering of results. Following the retrieval of data, the dataset had to be aggregated and collated into a structured machine-readable format for analysis. The dataset was cleaned to remove duplicates, website links and unfunded innovative projects. The literature shows that description of the project, pictures, and video footage are factors that increase consumers' desire which are associated with the backing of a project; hence, this study utilized the title, description, and year of projects for analysis [51]. A 10 per cent funding threshold was introduced to identify desirable projects. Through this process, we manually labelled and arrived at a final total of 8,021 projects labelled according to the project description, project amount reached, and year.

We used VOSviewer software, which applies an automatic text mining functionality through a linguistic filter that identifies noun phrases (a word sequence that consists of a noun and adjectives and ends with a noun) [66]. We utilized the collected data's project descriptions and years to study the results. Using the terms in the finalized dataset, we created a co-occurrence text-based map using a full counting method with a minimum number of at least three

occurrences. Following this, 2,194 terms met the threshold and 80 per cent of listed terms were selected. General terms such as "better way", "first time" and "full potential" were excluded to create and arrive at the distance-based term map which offer label, scatter and cluster density views. The distance reflects the strength and co-occurrence relationship between terms: the shorter the distance, the stronger the relationship. The co-occurrence matrix represents the co-occurrence of a term due its relevance leading to the closeness of terms to each other such that if a term has low co-occurrence across the data with other terms, the terms are positioned apart from one another. We visualized and clustered the data using centrality measures with the frequency of terms.

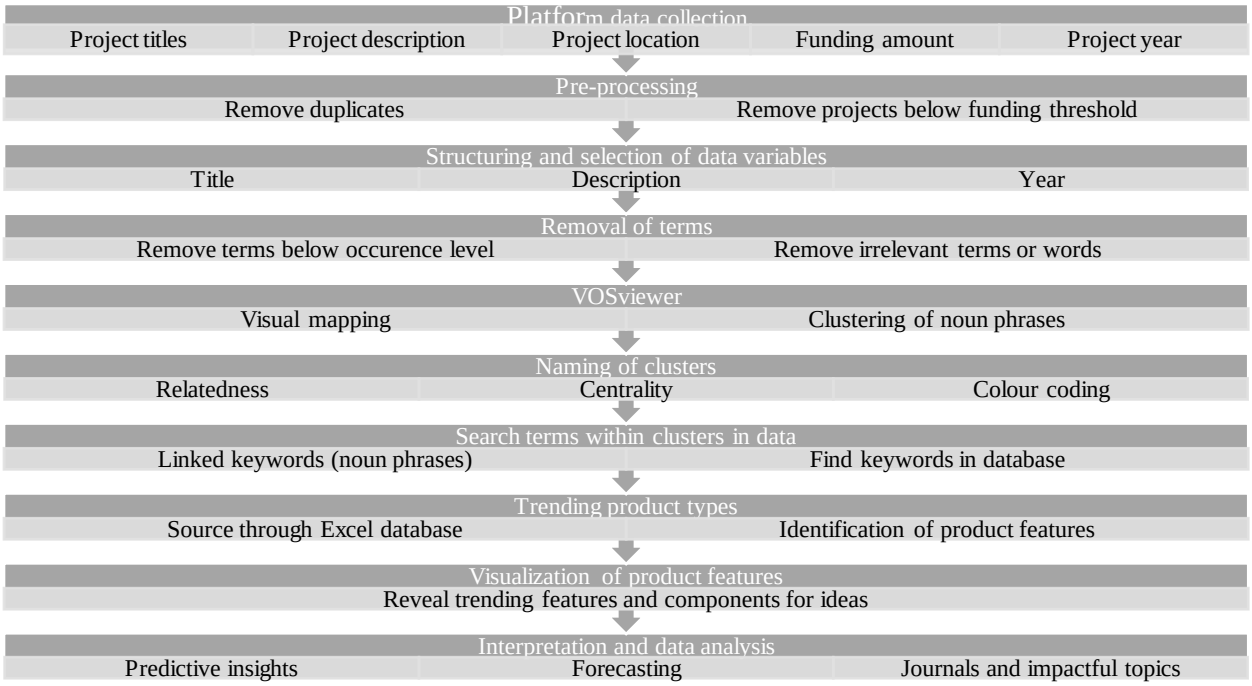

Fig. 2. Text mining process to identify product features from crowdfunding platforms

The retrieval of specific elements such as project title, project description, project links, project target goal, project amount reached, project category and year allowed us to examine funded projects. The VOSviewer software offers a multidimensional scaling (MDS) algorithm that has been utilized in previous studies [67, 68]. Creating a map with VOSviewer follows three stages

[69]: the similarity co-occurrence matrix, mapping of objects, and optimization to arrive at consistent results. As a result, a total of 313 unique keywords with a 2,087 total link strength and 1,038 links were identified for further analysis. Due to the number of technologies on crowdfunding platforms, we examine and cluster desired crowdfunded innovations on crowdfunding platforms using the term frequency-inverse document frequency (TD-IDF). We examine the present map based on the clustering association of terms in a co-occurrence matrix. The positioning of terms based on centrality calculations reveals the cluster and relevance of terms, as the VOSviewer software improves the visualization by colour coding each cluster and adjusting the size of terms based on centrality measures. The node size reveals the number of times a term appears with other terms within the data.

For interpretation, we examined each visual cluster by using the most to least occurring noun phrases in each cluster to locate products and assist with detecting product components and features in the database. Consequently, we classified, performed in-depth analysis, and interpreted our final visuals using qualitative analysis of a combination of terms within clusters and the database. To perform this step, we read through funded projects to confirm whether the terms identified are related to the cluster and reveal desirable product features. We selected products considering product descriptions. Afterwards, we labelled clusters accordingly. Furthermore, we chose each defined cluster and reintegrated the data into the software to arrive at key terms specific to the cluster, enabling the identification of product features. For example, the "electronic music accessories" cluster showed the term 'guitar' and using this term many products and product descriptions are located to reveal product features. To increase the robustness of the results, we utilized different VOSviewer visualization algorithms such as fractionalization and achieved the same clustering results. We examined the clusters in correlation to the purpose of each phase of the proposed theoretical framework. To determine the desirability of products, a 10 percent funding threshold was introduced to identify and scale

products according to three funding constructs: overfunded, mid-funded and low-funded. The identified products could then act as guiding opportunities for ideas during the pre-phase zero, phase zero, and phase one. To illustrate, we selected the 3D printing and application cluster and followed the step-by-step process to arrive at application areas based on funded targets to identify product opportunities. During the pre-phase, we examined the clustering visuals to arrive at focus areas by identifying dominant technologies. During phase zero, the focus areas were analysed based on existing technologies within clusters, project types and funding levels achieved. The existing technologies were then grouped into funding constructs. In phase one, the funding constructs enabled the identification of patterns whereby products and their underlying features were examined to aid opportunity identification. The validity and reliability of the applied approach was ensured by considering the following considerations: 1) The projects were respective of related clusters, 2) The labelling process done after the text mining analysis by the authors were completed by in-depth qualitative analysis of the dataset to agree on each cluster, 3) The funded projects were examined using single or a combinations of terms to understand whether the technologies are indicative of the appropriate cluster and 4) Ten percent project funding threshold were utilized to identify and examine projects.

IV. Findings and Analysis

This section provides the results of the text mining and classification method used to explore crowdfunding projects. We present the analysis of projects grouped into clusters and utilized to reveal opportunities for ideation. We then select a cluster as a case and propose identified opportunities.

A. Overview of Consumer Desired Innovations

The frequency of keywords can directly reflect the focus and type of innovation projects. We analysed the project titles and descriptions from 2009 to 2018 and identified the top 20

keywords based on co-occurrences. Figure 3 illustrates the top occurring terms, which directed the authors to further define and map the overarching technologies within each cluster.

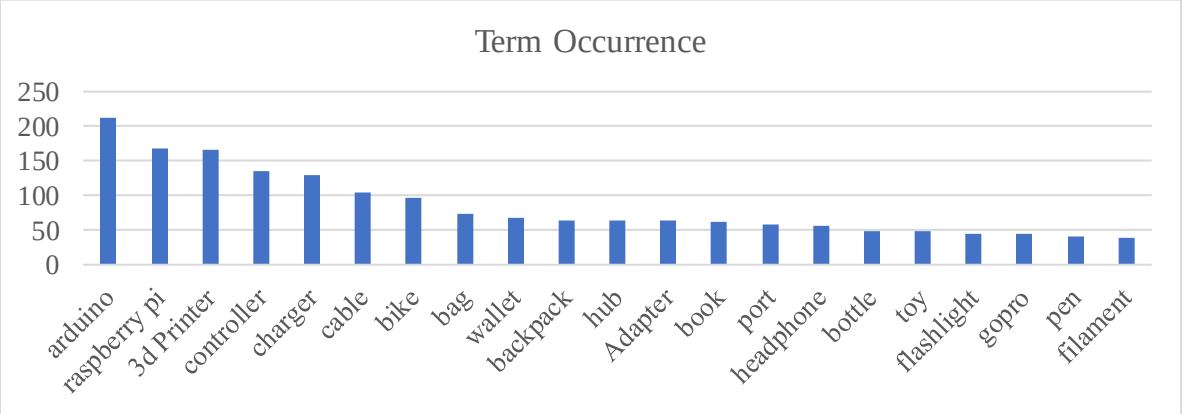

Fig. 3. Breakdown of clusters of consumer-desired innovations

Further analysis shows the gradual increase in desirable products in focus over time, which signifies how platforms can become a valuable source of seed capital for entrepreneurs and start-ups. The figure above reveals trending terms such as Arduino, Raspberry Pi, and 3D printer which entrepreneurs and start-ups are utilizing to develop prototypes, product concepts and innovations. For example, Arduino is utilized to create many incremental innovations in minor product improvements and product line extensions. 3D printing is seen to present a diverse range of innovative projects ranging from products new to the world, such as open source 3D printing and 3D printing filaments with materials such as hemp, to customized products such as 3D printer skull kits, 3D printed titanium rings, and a customizable speaker and sunrise alarm clock using 3D printing technology. The literature on the funding of crowdfunded projects reveals that consumers are attracted to more incremental innovation than radical innovation campaigns due to increased benefits and low learning costs. These incremental contributions can be utilized to provide signals of the wisdom of the crowd in terms of what people like and need within a product and its features.

B. Cluster Analysis for Consumer-Desired Innovations

The main aim of this section is to examine all the selected crowdfunding projects by applying clustering analysis and to link it to the FFE of the new product development and innovation process. Through this cleaning process, 313 terms were selected, resulting in nine cluster projects for analysis related to: A. home utilities (green); B. 3D printing and applications (orange); C. educational course and computing (red); D. electric device and robotics (pink); E. smart utilities (yellow); F. electronic music accessories (brown); G. outdoor accessories (turquoise); H. sports and travel (blue); and I. clothing and care (purple). Observing the top 20 keywords for each cluster, we can reveal and define the word map classified by VOSviewer in Figure 4. The colour mapping via VOSviewer for the clustering analysis represents the closeness and linkage between terms of desired innovative projects. The figure reveals trends in crowdfunding projects receiving better support and consumer interests. As illustrated, clusters which are similar in nature are close to each other. clusters B and C, which are technology-oriented clusters, are close to each other; likewise, cluster A and cluster I are home and clothing areas.

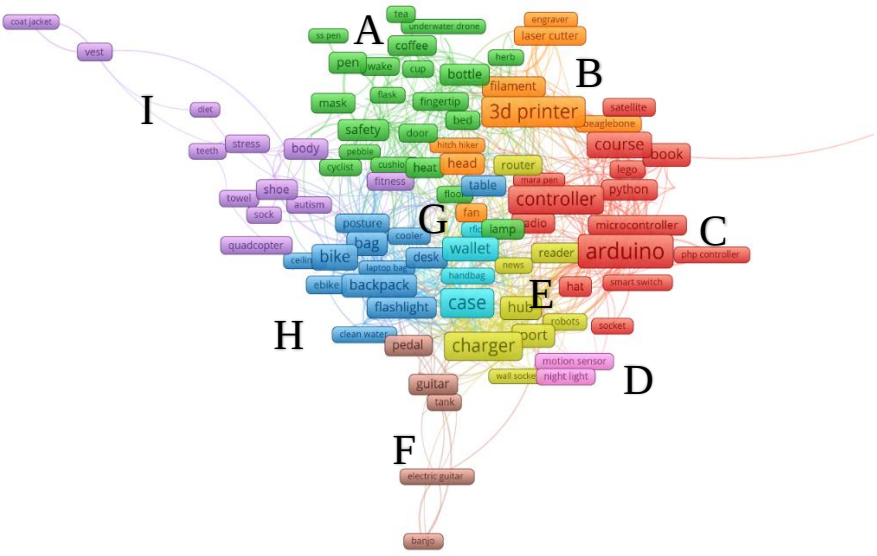

Fig. 4. Breakdown of clusters of consumer-desired innovations

This method enables us to reveal a pool of consumer-desired innovation within innovation clusters, thereby obtaining knowledge of what products interest consumers, which can further give a picture of what new innovative products to develop that would satisfy customers' needs in various markets. As prior research has paid attention to idea source, opportunity recognition, and customer involvement, and has identified the idea generation and selection process as an important stage, efforts have been made to curb ambiguity and lead to technology discovery.

We utilize noun phrases from clustered innovations to identify promising features from these desired products. To expand on the desired technologies as displayed in Table III, this study utilized noun phrases belonging to each cluster to identify trends in developed products, thereby guiding companies and entrepreneurs on which projects to focus on to develop incremental or radical innovations. Firstly, we examine the clusters ranging from A to I, and select the noun phrase based on the node size (term frequency). Many different product ideas can be identified from the database specific to the consumer's needs and requirements; however, we base our illustration on four specific clusters to show the process of arriving at trending products and features. For example, cluster G (outdoor accessories) has a noun phrase wallet, a product used by consumers to safeguard precious items. Wallet-related development trends reveal Bluetooth implementations such as the Bluetooth integrated wallets, where lost wallets can be found. Following the same process focusing on cluster E (smart utilities), we see applications such as portable solar USB chargers. Focusing on cluster A (home utilities), we see applications such as smart masks and a solar powered coffee roaster. Focusing on cluster I (clothing care), we see applications such as a rider-worn tail light brake light, a jacket designed with an air conditioning unit, etc. Focusing on cluster B (3D printing and applications), we see application areas such as 3D printers and a portable laser cutter. Focusing on H (sports and travel), we see application areas such as a folding ebike with a swappable

battery and smart-designed locks for secure homes. Table III reveals a list of consumer-desired innovations based on the cluster of desired innovations from crowdfunding platforms.

Table III. Consumer-desired innovations

Innovation clusters	Consumer-desired innovations
Home utilities (A)	Sleek design spotting machined pen; horse riding alert system; solar powered coffee roaster; boat hull underwater drone; self-heating jacket; smart mask; bionic wireless fish drone
3D printing and applications (B)	Desktop 3D printer; circuit board manufacturing equipment; laser engraver; portable laser cutter; portable laser engraver kit
Educational course and computing (C)	32-bit 48Mhz arm cortex processor; tablet powered smart mirror; miniature industrial robot arm; game adaptor; programmable spinner; digital pet robot; robotic arm; LED bow tie; graphic card shield; pocket drone
Electric device and robotics (D)	Small adjustable motion sensor night light; motion sensor built-in Wi-Fi hardware; hands-free smart electric skateboard; motion sensor basketballs
Smart utilities (E)	Slim phone charger with built-in cables; flexible audio jack; lightning MFi certified cable; photo and video enabled smart bracelet
Electronic music accessories (F)	Modular electric guitar and bass pickups; String Butler acrylic; electric guitar accessory; multi-tool guitar pedals; sound reactive LED mask
Outdoor accessories (G)	Bluetooth built-in leather wallet; solar powered phone charging wallet; case with strap; versatile phone case; Wi-Fi plug clock; Nixie tube; keychain backup battery; GPS tracking wallet (smart wallet)
Sports and travel (H)	Folding ebike with a swappable battery; multi travel bag; better back posture strap; 7-in-1 smart backpack; rechargeable multi-functional flashlight; smart side table with wireless charging; hand signal controlled bike light; e-bike kit; brightness filtering eyewear; unison flashing safety lights; smart designed locks for secure homes
Clothing and care (I)	Rider-worn tail light brake light; jacket designed with air conditioning unit; clothes deodorizing Boxx bag; range finding glove; touchscreen heating gloves; longest flying quadcopter

*Innovation significance based on 10 per cent threshold.

As shown above, using our approach, crowdfunding-related application areas can be quickly identified and examined (nine application areas are found). Each of these application areas can be examined and complemented by the funding accumulated to identify opportunities and derive ideas during the FFE phase.

Next, we will examine the 3D printing application area in depth to illustrate how such an approach can be integrated into the FFE phase. 3D printing is one of the best cases compared to other identified clusters as this technology is applied to develop many different innovations, and the technology itself is being developed with the support of the crowdfunding community. Thus, this cluster is selected for fruitful illustration of how companies can examine these developments to benefit the pre-phase, phase zero and phase one of the FFE phase.

C. Consumer-Desired Innovations: 3D Printing Ideas and Features

We applied our method based on the conceptual framework as shown in Figure 1 and illustrate the results in three phases as shown in Figure 5. Figure 5 illustrates the integrated FFE process complemented by crowdfunded data. This approach adds three further steps to other NPD models and innovation processes where innovation ideas are sourced from the open innovation environment. This generic model can be implemented in any field considering the firm's specific interests and capabilities. To illustrate this model, we selected the 3D printing and applications cluster. We elaborate on the steps and the underlying activities and practices to enable opportunity identification during the FFE phase of the innovation process.

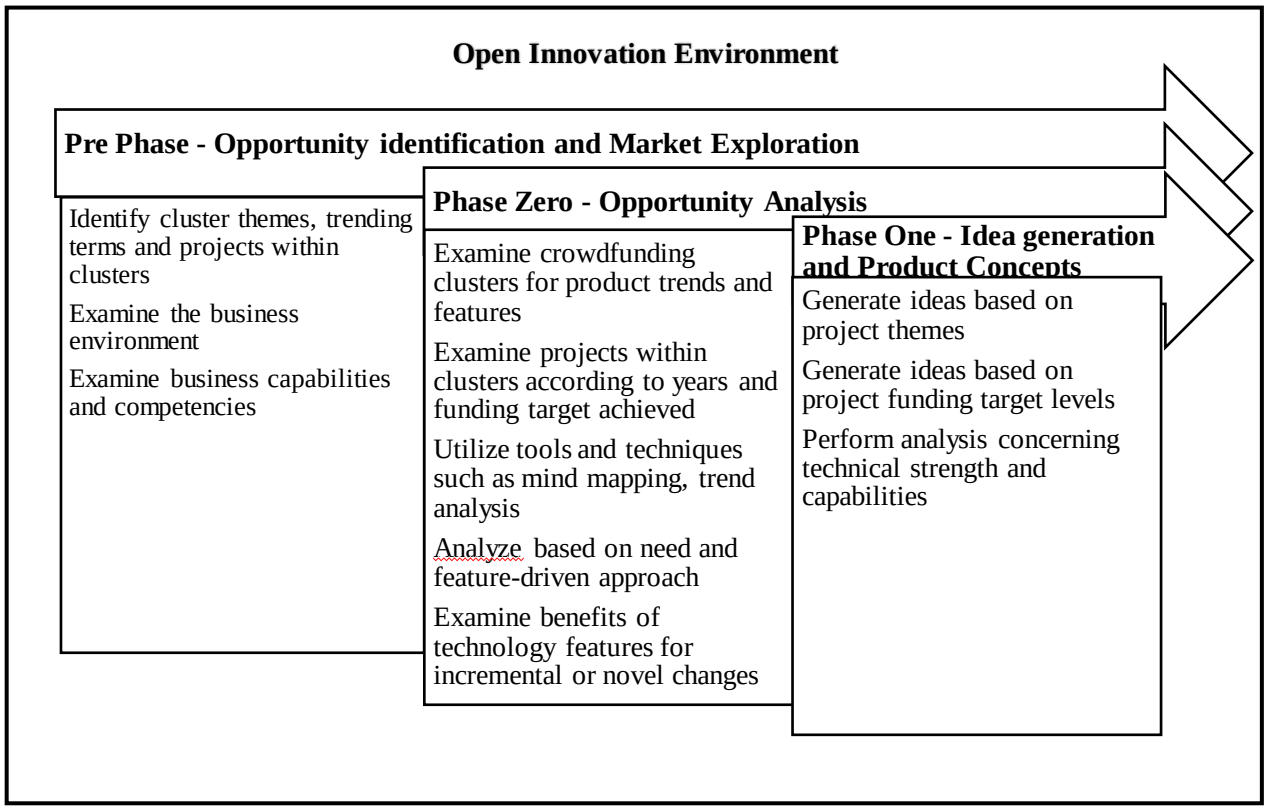

Fig. 5. Integrating crowdfunded data during the fuzzy front-end phase

1. Pre-Phase: Opportunity Identification and Market Exploration

During the pre-phase, examining one cluster to generate ideas relies on capitalizing on knowledge from crowdfunding data [40]. By applying information technology in this manner,

organizations can explore, perform market research, and gather competitive intelligence by identifying dominant technologies. A trend analysis exploring possible changes concerning the cluster arising from a sociocultural or technological trend enables potential opportunities to be arrived at. This would also rely on a clear understanding of an organization's strategic capabilities and competitive situation in the field. To understand the organization's competencies, organizations can investigate megatrends within the external and internal environments by analysing the political, social, technological, environmental and economic dimensions along with strengths, weaknesses, threats and opportunities to identify strategic areas that can be capitalized on.

1a. Crowdfunding Project: Identify Clusters

The breakdown of the 3D printing cluster reveals trending terms that are examined and combined with related terms in the same cluster or nearby cluster. Cluster B (3D printing and applications) is analysed and broken down into three subthemes: 1) user knowledge creativity and 3D printing development (C1); 2) desktop 3D printer, robotics and laser engraving (C2); and 3) open source 3D printing materials (C3). The high occurring terms are "student", "app", "pen", "desktop 3D printer", and "open source". The terms in each cluster enable us to locate projects funded by the crowd which can guide experts to identify trending innovative projects. Figure 6 illustrates the breakdown of the 3D cluster.

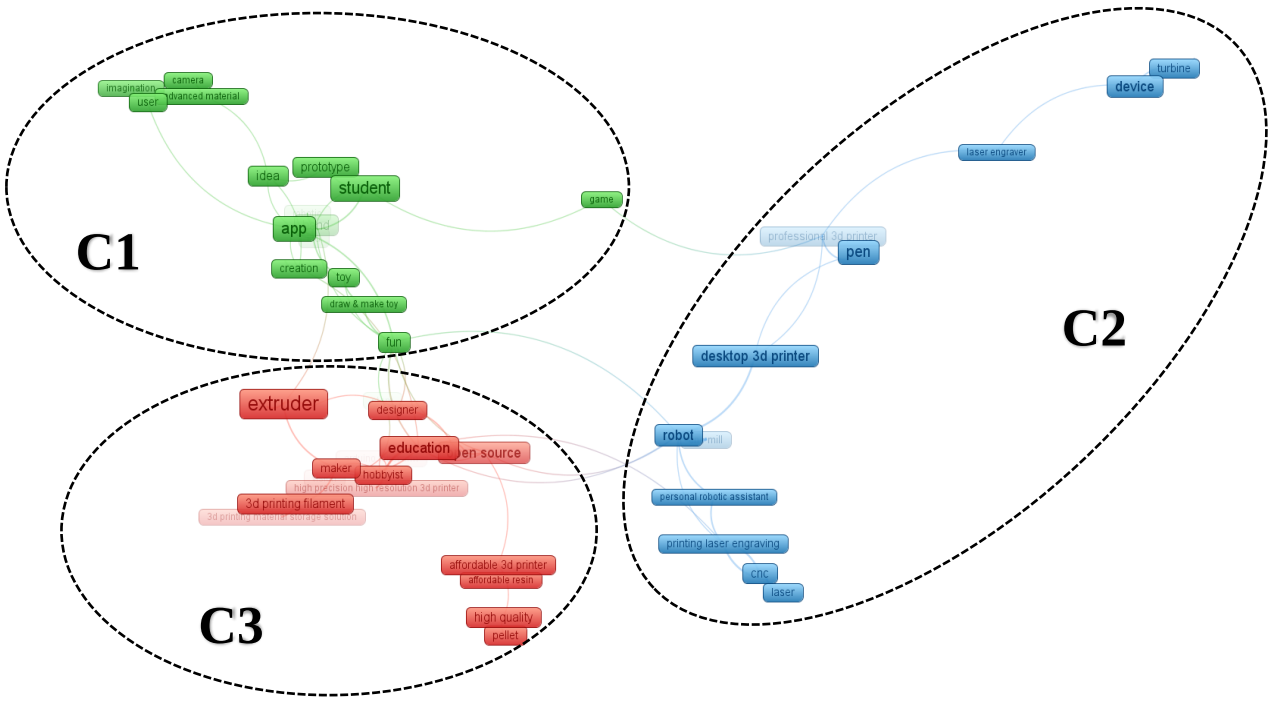

Fig. 6. Breakdown of 3D printing clusters

1b. Customer-Desired Projects: Examination

Based on the occurring terms and crowdfunded project percentage (CPP), we break the projects down into three-year groups (2009–2012, 2013–2015, 2016–2018) to identify projects that fall into three categories: overfunded, mid-funded and low-funded. Table IV illustrates ideas and product features based on the identified themes.

Table IV. Opportunities and product features

3D printing trends and potential investment areas	**2009–2012**	**CPP**	**2013–2015**	**CPP**	**2016–2018**	**CPP**
User knowledge creativity and 3D printing development (C1)	Sustainable 3D printing for blind students; Maxifab 3D printing framework	38% 488%	3D ThinkLink middle school 3D printer	113% 102%	Vulcan I rocket powered by 3D printed engine; EduExo 3D printable Arduino-powered kit	146% 271%
Desktop 3D printer, robotics and laser engraving (C2)	EasyMaker reconfigurable robot to make stuff; Dra-gun, a 3D copier/art precision tool	28% 17%	3D humanoid robotics; Titan 1 high res SLA 3D printer	121% 1374%	uArm Swift personal robotic assistant; Snapmaker all metal 3D printer	9569% 4554%
Open source 3D printing materials (C3)	The TangiBot Rapcraft rapid prototyping open source 3D printer	11% 27%	Portable and affordable 3D DLP printer; B9Creator open source photo-initiated polymer resin-based 3D printing system	24% 1027%	Eleven large open source affordable 3D printer; CowTech Ciclop open source 3D scanner	184% 1838%

Cluster 1 contains many examples concerning user knowledge creativity with 3D printing for early age school children, students, educational activities and creative learning in areas such as customizable pottery. Cluster 2 contains many examples concerning desktop 3D printing, robotics and laser engraving for personal and industrial manufacturing, portable laser engravers, robotic hand, mobile 3D printing units and producing prosthetic items. Cluster 3 contains many examples of open source 3D printing materials concerning expandable 3D printing systems, filament extruders, 3D printing kits and wireless 3D printers. Looking at year groups, we see a shift from incremental innovations such as all-metal automated printers and immobile printers to radical innovations where organizations produce mobile, networked portable printers, and a shift from plastic filaments to eco-friendlier materials such as hemp and ceramics for process applications. A mix between business and market drivers is observed due to the expected shift towards automation. The identified market drivers for the development and management of 3D printing stem from actions related to areas such as sustainability and the environment, evolving business models that integrate open processes (crowdsourcing, co-creation), individualization within the supply chain, the convergence of technologies due to digitization, and reduced cost and ease of developing customizable products. The business needs that fuel these changes range from the production of functional product parts to realignment of the supply chain, an increase in demand for personalized products, and a shift towards designing radical and complex products. Organizations will seek to emphasize efficiency in large-scale production of products and a reduced labour force in many industries. Examining the identified desired 3D printing clusters and funding levels, strategic areas and analysis by the R&D team can yield awareness of emerging technologies and a clear understanding of the competitive situation. Potential opportunities should be aligned with the overall business strategy, and it is thus important that an organization's leadership and culture supports market exploration [34].

2. Phase Zero: Opportunity Analysis

Phase zero relates to analysing the opportunities, positioning and benefits delivered by products. It is proposed that divergent thinking is essential for developing new ideas for products [15]. However, with the identified 3D printing cluster in view, organizations can define and generate concepts by analysing the product's technical features. Figure 5 shows the 3D printing themes, comprising a variety of projects with features that interest consumers along with some projects reaching overfunding levels, which signals stronger desire. These features can be studied and augmented based on creativity and knowhow to connect fields to achieve novel outcomes. The use of complementary tools and techniques such as brainstorming, trend analysis, problem solving techniques and mind mapping can help to identify opportunities, which can either be incremental changes to existing products or an entirely new product driven by the goal of the business [34, 40]. Business opportunities concern ideas such as customizable prosthetic development and production, customizable jewellery production, architectural models and engineering prototype and product development centres, personalized open 3D printing training centres, toys and electronics assembly production.

In order to guide the identification of opportunities, the funding levels of projects can be examined according to the target expected by project owners. The funding levels and utility signify the desire of the crowd for specific innovative projects. Furthermore, organizations should take a needs and feature-driven approach. A needs-driven approach can focus on the needs of current and future customers by analysing customer behaviour in existing markets and crowdfunding projects that fall within funding themes to search for opportunities. In addition, a feature-driven approach, exploring existing features within crowdfunding projects, can help spot new features and can result in new markets and customer behaviours. To support this process, methods such as the value proposition canvas are used to arrive at a strong product fit

addressing the most significant pains and gains from the customer profile [70]. Table V illustrates a breakdown of projects according to shifting crowd desires and funding levels.

Table V. Crowd desires and funding levels

User knowledge creativity and 3D printing development (C1)			
Years	Overfunded	Mid-funded	Low-funded
2009–2012	Maxifab 3D printing framework	PotteryPrint Imagine Create Fabricate, an app that enables kids to design their own unique works of pottery; sustainable 3D printing for blind students	Student-built lunar Rover prototype for Google Lunar X PRIZE
2013–2015	3D ThinkLink middle school 3D printer	Modular metal construction system for building computerized machines	RoboGut, a talking singing, dancing, educational robot project kit for kids
2016–2018	Vulcan I rocket powered by 3D printed engine; EduExo 3D printable Arduino-powered kit		
Desktop 3D printer, robotics and laser engraving (C2)			
2009–2012		EasyMaker, a reconfigurable robot to make stuff	Enter the Dra-gun, a 3D copier/art precision tool
2013–2015	3D humanoid robotics Titan 1 high res SLA 3D printer		MAKI, the 3D printable friendly humanoid robot
2016–2018	uArm Swift, a personal robotic assistant; Snapmaker, all metal 3D printer	ShopArm, a 3D printed Arduino-based robot arm for repetitive jobs	Closed loop industrial 3D printer based on open source software
Open source 3D printing materials (C3)			
2009–2012		Rapcraft, rapid prototyping open source 3D printer	PrusAL modular opensource 3D printer
2013–2015	B9Creator, an open source photo-initiated polymer resin-based 3D printing system	Thingystock, an expandable build volume open Delta 3D printer	PrismX, an open source 3D printer
2016–2018	Eleven 3D printer, a large open source affordable 3D printer CowTech Ciclop, an open source 3D scanner	Open source feature-rich networked touchscreen Gigabot 3D printer	Evezor, an open source robotic arm manufacturing platform

3. Phase One

Phase one relates to idea generation, concept specification and technical feasibility. In this phase, the modification and combining of elements into new ones enables the generation of ideas to be applied to new problems, industries or markets [31, 35]. The 3D printing cluster in Table III can be analysed according to years and funding levels, which range between overfunded, mid-funded and low-funded themes. Examining projects within each theme can fuel the idea generation process; communication with users familiar with projects on crowdfunding platforms and existing consumers of the organization, and collaboration with existing entrepreneurs and companies, can enhance this process. The proof of concept details can be established based on analysing the technical features that could aid authorization of full-

scale development and the allocation of resources. Furthermore, based on the organization's SWOT analysis to understand its own technical strengths and capabilities, the selection of an idea concept can be derived. Due to the need to decide which concept to focus on, the ideas team can utilize a voting mechanism to reach a consensus. The ideas team can complement examining the 3D printing clusters with needs and feature-driven research to identify actionable possibilities for a product concept. The ideas team can recombine knowledge derived from the previous phases and apply it in a new setting. With guidance from identified clusters and the strategic capabilities of the organization, the goal of arriving at suitable product concepts can be validated from a customer desirability, technical feasibility and economic viability perspective [34]. The selected product concepts can then be developed further to be brought into the market. Moving on to the next stage of the innovation process, organizations can test business concepts by examining the dimensions of business risk and attractiveness to filter out potential concepts and define potential customers. In addition, market segmentation can be utilized to divide customers into different attributes to link product concepts with the market and target them [71].

Generally, following a broader look at 3D printing with crowd desires, contributions towards projects, and changing market needs, it is proposed that technology is utilized in industries such as fashion, jewellery, health (medical or dental), automotive, construction, consumer products, electronics, aerospace, tools and moulding materials for efficiency in production and sustainability [72]. However, these results may not apply to some sectors. From this perspective, crowdfunders can generally be considered innovation enthusiasts. Currently, companies rely on market trends and customer feedback to develop radical or incremental innovations. With the growing importance of satisfying consumer needs, the above-mentioned process can help entrepreneurs and companies detect what features or components can be inbuilt during the development of new products. These features can be detected and selected

based on the company's technological capabilities and developed to satisfy the market. Uncovering a clustered list of consumer-desired innovations can assist designers, developers and market leaders during the FFE phase.

V. Discussion

The results presented in this study on crowdfunding platforms can help resolve the misconception that data obtained from consumer-oriented opportunity sources are of limited and low quality in guiding organizations in new product and service development. Based on the application and development of opportunity-led FFE methods, we aim to contribute to the literature from a theoretical and methodological perspective.

Previous authors have proposed many prevalent NPD processes, and the activities of the front end of NPD can either be in three stages or a five-gate process [34]. However, we propose an altered opportunity-led process complemented by the involvement of the open community. We have illustrated how the opportunity-led process with a customer-centric approach works in an open innovation environment (see Figure 5). We have explained in detail what type of considerations need to be followed in each step, starting from the initial cluster of themes to the final step where ideas are selected. We contribute to the literature by asserting that opportunity-led new product development processes based on examining crowdfunding projects can help seize opportunities, complemented by effective techniques to transform identified trending product features and projects in correlation with funding levels into ideas for successful innovative performance. Our approach to the crowdfunding platform was as an opportunity source, whereby we classified technologies into nine categories and cluster projects to gain insights and identify opportunities, ideas and product features during the FFE phase of the innovation process [73]. Afterwards, we implemented this approach in respect of 3D printing technologies to illustrate how the proposed process would work. We contribute to

the literature by revealing how organizations and entrepreneurs can effectively use crowdfunding platforms and data analytics in the form of funding levels to uncover patterns and identify and exploit opportunities [63]. Kutvonen and Torkkeli [8] also proposed the need for more studies promoting the adoption of an open innovation viewpoint in the FFE phase of the innovation process. In contrast to Seo et al. [61] and Ozcan et al. [12], we propose the use of crowdfunding platforms as a means for external concept exploitation by enabling opportunity identification and supporting idea generation and collaboration with companies.

Previous authors examined using patents, social media and scientific publications to identify opportunities and ideas [12, 55]. However, these studies utilized industry-related data sources, complemented heavily by forecasting methods, that might not necessarily relate to consumers' desires. Furthermore, the use of techniques such as customer surveys and focus groups can be expensive and less duplicatable. We have uncovered that innovative products emerge from crowdfunding campaigns based on collaboration with consumers. Therefore, we propose crowdfunding as a fruitful medium to acquire information that enables the identification of opportunities that can help existing firms develop new products and innovations. We make a practical contribution by extending text mining techniques to a novel data source that encompasses information on products desired from both industrial and consumer perspectives. By examining the closeness of text in the visualized output, we identified trending products and features from either the same domain or cross-domain. This can assist in decision-making at the ideation stage of an NPD process.

Considering the theoretical contribution, we propose that organizations can benefit from the outside-in approach, whereby organizations can open up innovation processes to external inputs and contributions [36]. However, the benefits from outside-in exploitation need to be aligned and supported by factors such as the structure, leadership and culture, which are believed to influence an organization’s innovation [74, 75, 76]. On the other hand, the

management of technology development can benefit from mediums such as crowdfunding platforms as a source for opportunity identification. According to Chesbrough [77] and Teece [78], the management of technology is a key component of organizational strategy and competitive advantage. Lack of information is also a limiting factor for the front end. Utilizing complementary tools and techniques to support the innovation process in an era of digitalization can enable NPD teams to arrive at more successful innovations. Hence, this study adds value to the field and t previous studies with an open innovation-based NPD and innovation processes.

In correlation with studies such as Lee and Sohn [28], we propose that the crowd can be co-creators of value, whereby organizations can detect desirable product features from crowdfunding platforms, analyze through organization routines to identify opportunities and transform these inputs into revenue-producing products and services. We contribute by revealing that crowdfunding platforms can be a medium for information beyond finance to gain insights that can enrich an organization's product development [23]. As designers, product and innovation managers need tools and processes for the ideation phase to identify product opportunities: our unique framework enables the rapid retrieval of ideas based on the desires of current and potential consumers, which are more validated and user-centric and less risky. Start-ups' acceptance of utilizing crowdfunding platforms and the crowd represents a disruptive change in how new ventures are funded. It also serves as a fruitful way of testing product concepts and extending existing theories.

VI. Conclusion

To our knowledge, this is one of the few studies that illustrate how opportunities can be identified through text mining crowdfunding data for integration into the FFE phase. Our study indicates that crowdfunding has substantial potential to facilitate innovation through new

funding sources and crowd participation. However, crowdfunding is still a rapidly growing phenomenon. The output of this study reveals how the activities of consumers can be a significant signal for organizations to determine the desires of consumers. The pledging of funds provides evidence regarding the generation and selection of ideas that can help the organization arrive at incremental ideas during the FFE phase of the NPD process.

This study suggests using open channels such as crowdfunding platforms to aid awareness of market changes and consumer desires. The utilization of divergent information which is structured can enable organizations to develop stronger relationships with their customer base and market as a whole. Most studies, such as Wang and Chen [11], examine patents to reveal technology opportunities. However, this study contributes to the literature by proposing how crowdfunding data sources can be utilized as a guide during the FFE phase of the innovation process. Our proposed system clusters innovative projects and identifies popular consumer-desired innovations specific to developing innovation opportunities. This paper correlates with studies such as Ozcan et al. [12] and Lee et al. [28], which propose that consumer innovation-related data can be retrieved and linked to product development.

Another contribution of this study is identifying how crowdfunding approaches foster innovation. Crowdfunding is a medium to source funds and acquire feedback for further development. Furthermore, it was identified that projects require funding from the ideation stage and the development, prototyping and marketing stages of the innovation process to ease commercialization. It is evident that crowdfunding can be manifested during each stage of the innovation process and can positively influence innovation performance. In summary, divergence in thinking and information is necessary for organizations to utilize this approach.

There are limitations to this study. There is limited evidence of utilizing crowdfunding platforms as a data source. In addition, although we have examined the 3D printing cluster in

this study, our methods have not yet been applied to other identified clusters or fields. Overall, the literature can benefit from complementing this process with a qualitative approach by engaging with product innovation managers. Future studies can further examine platform functionalities such as reviews and comments sections to reveal the benefits and challenges of utilizing these functionalities. This can guide the development of new products that can disrupt markets, thereby increasing organization profit margins. Future studies can improve our approach by introducing a qualitative assessment of the selection method considering the companies' capabilities or future innovation portfolios. In summary, as the characteristics of the front-end phase are considered informal, relying on seeking knowledge, experimenting and creativity, this approach can act as a compass for fulfilling the crowd's unknown desires.

References

[1] B. Schemmann, A. M. Herrmann, M. M. H. Chappin, and G. J. Heimeriks, "Crowdsourcing ideas: Involving ordinary users in the ideation phase of new product development," *Research Policy*, vol. 45, no. 6, pp. 1145–1154, 2016.

[2] D. A. Norman and R. Verganti, "Incremental and radical innovation: Design research vs. technology and meaning change," *Design Issues*, vol. 30, no. 1, pp. 78–96, 2014.

[3] K. Eling and C. Herstatt, "Managing the front end of innovation—Less fuzzy, yet still not fully understood," *Journal of Product Innovation Management*, vol. 34, no. 6, pp. 864–874, 2017.

[4] I. Verheul and L. Van Mil, "What determines the growth ambition of Dutch early-stage entrepreneurs?" *International Journal of Entrepreneurial Venturing*, vol. 3, no. 2, 2011.

[5] S. E. Reid, U. de Brentani, and E. J. Kleinschmidt, "Divergent thinking and market visioning competence: An early front-end radical innovation success typology," *Industrial Marketing Management*, vol. 43, no. 8, pp. 1351–1361, 2014.

[6] H. Moon and S. H. Han, "A creative idea generation methodology by future envisioning from the user experience perspective," *International Journal of Industrial Ergonomics*, vol. 56, pp. 84–96, 2016.

[7] E. Loanzon, J. Provenzola, B. Siriwannangkul, and M. Al Mallak, "Netnography: Evolution, trends, and implications as a fuzzy front end tool," in *2013 Proceedings of PICMET' 13: Technology Management in the IT-Driven Services (PICMET)*, IEEE, pp. 1572–1593, 2013.

[8] A. Kutvonen, and M. T. Torkkeli, "Opening the fuzzy front end: A synthesis of two theories. In *PICMET'09-2009 Portland International Conference on Management of Engineering & Technology*, IEEE, pp. 2029-2036, 2009.

[9] G. Kim and J. Bae, "A novel approach to forecast promising technology through patent analysis," *Technological Forecasting and Social Change*, vol. 117, pp. 228–237, 2017.

[10] H. Zhang, H. Rao, and J. Feng, "Product innovation based on online review data mining: A case study of Huawei phones," *Electronic Commerce Research*, vol. 18, no. 1, pp. 3–22, 2018.

[11] J. Wang and Y. J. Chen, "A novelty detection patent mining approach for analyzing technological opportunities," *Advanced Engineering Informatics*, vol. 42, p. 100941, 2019.

[12] S. Ozcan, M. Suloglu, C. O. Sakar, and S. Chatufale, "Social media mining for ideation: Identification of sustainable solutions and opinions," *Technovation*, vol. 107, p. 102322, 2021.

[13] K. Wyrtki, M. Röglinger, and M. Rosemann, "Opportunity-led ideation: How to convert corporate opportunities into innovative ideas," *Creativity and Innovation Management*, vol. 30, no. 3, pp. 523–541, 2021.

[14] J. Frishammar, E. Dahlskog, C. Krumlinde, and K. Yazgan, "The front end of radical innovation: A case study of idea and concept development at prime group," *Creativity and Innovation Management*, vol. 25, no. 2, pp. 179–198, 2016.

[15] J. Nicholas, A. Ledwith, and J. Bessant, "Selecting early-stage ideas for radical innovation: Tools and structures," *Research-Technology Management*, vol. 58, no. 4, pp. 36–44, 2015.

[16] J. Van den Ende, L. Frederiksen, and A. Prencipe, "The front end of innovation: Organizing search for ideas," *Journal of Product Innovation Management*, vol. 32, no. 4, pp. 482–487, 2015.

[17] L. Broennum and C. Clausen, "Enabling front end of innovation in a mature development company," in *Proceedings of the 20th International Conference on Engineering Design (ICED15): Innovation and Creativity, Design Society*, pp. 235–244, 2015.

[18] S. Achiche, F. P. Appio, T. C. McAloone, and A. Di Minin, "Fuzzy decision support for tools selection in the core front end activities of new product development," *Research in Engineering Design*, vol. 24, pp. 1–18, 2013.

[19] Y. Geum and Y. Park, "How to generate creative ideas for innovation: A hybrid approach of WordNet and morphological analysis," *Technological Forecasting and Social Change*, vol. 111, pp. 176–187, 2016.

[20] Q. Zhang and W. J. Doll, "The fuzzy front end and success of new product development: A causal model," *European Journal of Innovation Management*, vol. 4, no. 2, pp. 95–112, 2001.

[21] S. Ozcan, D. Boye, J. Arsenyan, and P. Trott, "A scientometric exploration of crowdsourcing: Research clusters and applications," *IEEE Transactions on Engineering Management*, vol. 69, no. 6, 2020.

[22] V. Butticè and E. Ughetto, "What, where, who, and how? A bibliometric study of crowdfunding research," *IEEE Transactions on Engineering Management, IEEE,* Jan 26, 2021.

[23] J. V. Da Cruz, "Beyond financing: Crowdfunding as an informational mechanism," *Journal of Business Venturing*, vol. 33, no. 3, pp. 371–393, 2018.

[24] E. Mollick and A. Robb, "Democratizing innovation and capital access: The role of crowdfunding," *California Management Review*, vol. 58, no. 2, pp. 72–87, 2016.

[25] V. Kayser and K. Blind, "Extending the knowledge base of foresight: The contribution of text mining," *Technological Forecasting and Social Change*, vol. 116, pp. 208–215, 2017.

[26] I. Spitsberg, M. J. Verti, S. Brahmandam, and G. W. Coulston, "Capitalizing on emerging technologies: A path to creating opportunities in new markets," *Research-Technology Management*, vol. 58, no. 4, pp. 17–27, 2015.

[27] K. Christensen, S. Nørskov, L. Frederiksen, and J. Scholderer, "In search of new product ideas: Identifying ideas in online communities by machine learning and text mining," *Creativity and Innovation Management*, vol. 26, no. 1, pp. 17–30, 2017.

[28] W. S. Lee and S. Y. Sohn, "Discovering emerging business ideas based on crowdfunded software projects," *Decision Support Systems*, vol. 116, pp. 102–113, 2019.

[29] B. Khademi, H. Lampela, and K. X. Smyrnios, "A roadmap for systematically identifying opportunities in ecosystems using scientific publications data," *Technology Innovation Management Review*, vol. 11, no. 1, 2021.

[30] D. Ahlstrom, "Innovation and growth: How business contributes to society," *Academy of Management Perspectives*, vol. 24, no. 3, pp. 11–24, 2010.

[31] J. Yoon, H. Park, W. Seo, J.-M. Lee, B. Coh, and J. Kim, "Technology opportunity discovery (TOD) from existing technologies and products: A function-based TOD framework," *Technological Forecasting and Social Change*, vol. 100, pp. 153–167, 2015.

[32] H. Chesbrough, "From open science to open innovation," *Institute for Innovation and Knowledge Management, ESADE*, Aug 12, 2015.

[33] H. Lee, K. Choi, D. Yoo, Y. Suh, S. Lee, and G. He, "Recommending valuable ideas in an open innovation community: A text mining approach to information overload problem," *Industrial Management & Data Systems*, vol. 118, no. 4, 2018.

[34] P. Koen, G. Ajamian, R. Burkart, A. Clamen, J. Davidson, R. D'Amore, C. Elkins, K. Herald, M. Incorvia, A. Johnson, R. Kartol, R. Seibert, A. Slavejkov, and K. Wagner, "Providing clarity and a common language to the 'fuzzy front end'," *Research-Technology Management*, vol. 44, no. 2, pp. 46–55, 2001.

[35] R. Cooper and S. J. Edgett, "Ideation for product innovation; What are the best methods?" *Reference Paper 29, Product Innovation Best Practices Series*, Product Development Institute, 2008.

[36] H. Chesbrough, "Managing open innovation," *Research-Technology Management*, vol. 47, no. 1, pp. 23–26, 2004.

[37] J. K. Christiansen and M. Gasparin, "Managing controversies in the fuzzy front end," *Creativity and Innovation Management*, vol. 25, no. 4, pp. 500–514, 2016.

[38] A. Riel, M. Neumann, and S. Tichkiewitch, "Structuring the early fuzzy front-end to manage ideation for new product development," *CIRP Annals*, vol. 62, no. 1, pp. 107–110, 2013.

[39] H. Liu, J. Goulding, and T. Brailsford, "Towards computation of novel ideas from corpora of scientific text," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Springer, 2015, pp. 541–556.

[40] P. Koen, G. Ajamian, S. Boyce, A. Clamen, S. Fountoulakis, A. Johnson, P. Puri, and R. Seibert, "Fuzzy front end: Effective methods, tools, and techniques," *The PDMA Toolbook 1 for New Product Development*, 2002.

[41] S. A. Alvarez and J. B. Barney, "Discovery and creation: Alternative theories of entrepreneurial action," *Strategic Entrepreneurship Journal*, vol. 1, no. 1–2, pp. 11–26, 2007.

[42] D. J. Teece, "Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance," *Strategic Management Journal*, vol. 28, no. 13, pp. 1319–1350, 2007.

[43] J. Dahlqvist and J. Wiklund, "Measuring the market newness of new ventures," *Journal of Business Venturing*, vol. 27, no. 2, pp. 185–196, 2012.

[44] P. Davidsson, "Entrepreneurial opportunities and the entrepreneurship nexus: A re-conceptualization," *Journal of Business Venturing*, vol. 30, no. 5, pp. 674–695, 2015.

[45] J. S. McMullen, L. A. Plummer, and Z. J. Acs, "What is an entrepreneurial opportunity?" *Small Business Economics*, vol. 28, no. 4, pp. 273–283, 2007.

[46] H. Moon, S. H. Han, and J. Kwahk, "A MORF-Vision method for strategic creation of IoT solution opportunities," *International Journal of Human–Computer Interaction*, vol. 35, no. 10, pp. 821–830, 2019.

[47] P. A. Koen, H. M. J. Bertels, and E. Kleinschmidt, "Managing the front end of innovation. Part I: Results from a three-year study," *Research-Technology Management*, vol. 57, no. 2, pp. 34–43, 2014.

[48] C. Simms and P. Trott, "Packaging development: A conceptual framework for identifying new product opportunities," *Marketing Theory*, vol. 10, no. 4, pp. 397–415, 2010.

[49] L. Dahlander and H. Piezunka, "Open to suggestions: How organizations elicit suggestions through proactive and reactive attention," *Research Policy*, vol. 43, no. 5, pp. 812–827, 2014.

[50] C. A. de Mattos, K. O. Kissimoto, and F. J. B. Laurindo, "The role of information technology for building virtual environments to integrate crowdsourcing mechanisms into the open innovation process," *Technological Forecasting and Social Change*, vol. 129, pp. 143–153, 2018.

[51] C. S. R. Chan and A. Parhankangas, "Crowdfunding innovative ideas: How incremental and radical innovativeness influence funding outcomes," *Entrepreneurship Theory and Practice*, vol. 41, no. 2, pp. 237–263, 2017.

[52] M. A. Stanko and D. H. Henard, "Toward a better understanding of crowdfunding, openness and the consequences for innovation," *Research Policy*, vol. 46, no. 4, pp. 784–798, 2017.

[53] V. Kuppuswamy and B. L. Bayus, "Crowdfunding creative ideas: The dynamics of project backers," in *The Economics of Crowdfunding*, Springer, 2018, pp. 151–182.

[54] S. Lee, B. Yoon, and Y. Park, "An approach to discovering new technology opportunities: Keyword-based patent map approach," *Technovation*, vol. 29, no. 6–7, pp. 481–497, 2009.

[55] C. Kim and H. Lee, "A patent-based approach for the identification of technology-based service opportunities," *Computers & Industrial Engineering*, vol. 144, p. 106464, 2020.

[56] R. Kohli and N. P. Melville, "Digital innovation: A review and synthesis," *Information Systems Journal*, vol. 29, no. 1, pp. 200–223, 2019.

[57] J. Guo, P. Jiang, J. W. Guo, and R. H. Tan, "Research of idea generation process for fuzzy front end based on patent analysis," in *2012 IEEE International Conference on Industrial Engineering and Engineering Management*, IEEE, pp. 1190–1194, Dec, 2012.

[58] C. Tucker and H. Kim, "Predicting emerging product design trend by mining publicly [1] available customer review data," in *DS 68-6: Proceedings of the 18th International Conference on Engineering Design (ICED 11), Impacting Society through Engineering Design, Vol. 6: Design Information and Knowledge*, Lyngby/Copenhagen, Denmark, 15–19 Aug, 2011.

[59] D. Zhu and A. L. Porter, "Automated extraction and visualization of information for technological intelligence and forecasting," *Technological Forecasting and Social Change*, vol. 69, no. 5, pp. 495–506, 2002.

[60] K. Laursen and A. Salter, "Open for innovation: The role of openness in explaining innovation performance among UK manufacturing firms," *Strategic Management Journal*, vol. 27, no. 2, pp. 131–150, 2006.

[61] W. Seo, J. Yoon, H. Park, B. Coh, J.-M. Lee, and O.-J. Kwon, "Product opportunity identification based on internal capabilities using text mining and association rule mining," *Technological Forecasting and Social Change*, vol. 105, pp. 94–104, 2016.

[62] P. Trott, *Innovation Management and New Product Development*, 6th ed., Pearson, 2017.

[63] J.-M. Sahut, L. Iandoli, and F. Teulon, "The age of digital entrepreneurship," *Small Business Economics*, vol. 56, pp. 1159–1169, 2021.

[64] C. Herstatt and B. Verworn, "The 'fuzzy front end' of innovation," in *Bringing Technology and Innovation into the Boardroom*. Springer, 2004, pp. 347–372.

[65] M. Thelwall and D. Stuart, "Web crawling ethics revisited: Cost, privacy, and denial of service," *Journal of the American Society for Information Science and Technology*, vol. 57, no. 13, pp. 1771–1779, 2006.

[66] K. Demeter, L. Szász, and A. Kő, "A text mining based overview of inventory research in the ISIR special issues 1994–2016," *International Journal of Production Economics*, vol. 209, pp. 134–146, 2019.

[67] M. Sedighi, "Application of word co-occurrence analysis method in mapping of the scientific fields (case study: the field of Informetrics)," *Library Review*, vol. 65, no. 1/2, pp. 52–64, 2016.

[68] A. S. Rusydiana, S. Rahardjo, and W. S. Soeparno, "Mapping research on halal logistics using VOSviewer," *Library Philosophy and Practice*, vol. 5895, 2021.

[69] A. Papachroni, L. Heracleous, and S. Paroutis, "In pursuit of ambidexterity: Managerial reactions to innovation–efficiency tensions," *Human Relations*, vol. 69, no. 9, pp. 1791–1822, 2016.

[70] A. Osterwalder, Y. Pigneur, G. Bernarda, G., and A. Smith, *Value Proposition Design: How to Create Products and Services Customers Want*. John Wiley & Sons, 2015.

[71] G.T. Jou and B. J. C. Yuan, "Utilizing a novel approach at the fuzzy front-end of new product development: A case study in a flexible fabric supercapacitor," *Sustainability*, vol. 8, no. 8, p. 740, 2016.

[72] D. A. Arcos-Novillo and D. Güemes-Castorena, "Development of an additive manufacturing technology scenario for opportunity identification: The case of Mexico," *Futures*, vol. 90, pp. 1–15, 2017.

[73] O. E. Kurt, C. Ucler, and O. Vayvay, "From ideation towards innovation: Pillars of front-end in new product development," *Pressacademia Procedia*, vol. 5, no. 1, pp. 71–79, 2017.

[74] H. Van de Ven and M. S. Poole, "Explaining development and change in organizations," *Academy of Management Review*, vol. 20, no. 3, pp. 510–540, 1995.

[75] J. Tidd and J. R. Bessant, *Managing Innovation: Integrating Technological, Market and Organizational Change*. John Wiley & Sons, Nov, 2020.

[76] K. M. Eisenhardt and J. A. Martin, "Dynamic capabilities: What are they?" *Strategic Management Journal*, vol. 21, no. 10–11, pp. 1105–1121, 2000.

[77] H. Chesbrough, "Business model innovation: Opportunities and barriers," *Long Range Planning*, vol. 43, no. 2–3, pp. 354–363, 2010.

[78] D. J. Teece, "Dynamic capabilities and entrepreneurial management in large organizations: Toward a theory of the (entrepreneurial) firm," *European Economic Review*, vol. 86, pp. 202–216, 2016.