Article

Exploring Technological Trends in Logistics: Topic Modeling-Based Patent Analysis

Donghyun Choi and Bomi Song *

School of Air Transportation and Logistics, Korea Aerospace University, Goyang-si, Gyeonggi-do 10540, Korea; dchoi@kau.ac.kr

* Correspondence: bmsong@kau.ac.kr; Tel.: +82-2-300-0347

check for updates

Abstract: With the strategic importance of discerning opportunities and threats from technological development to achieve sustainable competitiveness, exploring technological trends becomes critical for a successful technology strategy in logistics. Given the rapid pace of development and varying technological options, logistics also increasingly requires methodological support and appropriate data to reduce the complexity and burden of exploring technology trends. While previous research has largely relied on experts' insights, the value of patent-based approaches for exploring technological trends has been underestimated in logistics. To address this gap, this study proposes a topic modeling-based approach using logistics-related patents registered at the United States Patents and Trademark Office (USPTO). The core of the suggested approach is latent Dirichlet allocation (LDA), allowing the identification of logistics-related technological topics behind patents. The topics identified by LDA are further investigated regarding both filed-level and firm-level trends. The suggested approach is expected to offer implications of the use of patents for the purpose of exploring the trends of technology development outside the organization in logistics. In addition, we believe that the information on the technological topics and their trends generated by the suggested approach can offer an enhanced understanding of the technological landscape in logistics.

Keywords: technological trends in logistics; technology trend exploration; patent analysis; topic modeling; latent dirichlet allocation

1. Introduction

Due to the easing of trade barriers and the introduction of developed technologies, logistics has become a strategical concern and a key for companies to develop sustainable competitive advantages [1–3]. One of the critical factors making logistics a source of differentiation is technology [4,5], and this role is gaining more significance and diversity with recent developments in technology [6]. It is well-recognized that the appropriate use of technologies can enhance the core function of logistics [7,8], as well as create new business models in logistics [9,10]. For instance, identification and data capture technologies such as bar-coding and radio frequency identification (RFID) have been widely used to track items and collect real-time data across the whole supply chain [11]. Those data can also be used with information systems, such as point-of-sales systems [11] and warehouse management systems [12], for effective management and decision making. Furthermore, leading companies like Amazon use these data to identify demand patterns and implement anticipatory shipping [13]. These technologies can also reduce the uncertainty involved in complex supply chains and fast-changing markets [14]. Thus, increasing and varying technological options for innovation may result in more opportunities for logistics firms to use advanced technologies for their business [4,5,15–17]. Accordingly, technological capability has become one of the strongest guarantors for survival in fast-changing, competitive environments by seeking sustainable competitiveness [9,18,19].

At the same time, these technology-driven changes present logistics firms with great challenges to overcome [20]. Technology is transforming value chains and expanding and blurring the area of logistics, thereby breaking-down traditional industry boundaries [1]. Established players must compete with new, unexpected competitors, which are often armed with the latest technologies and compelling business models [18]. Unlike the traditional players, the new players are not mere technology adopters; rather, they are more likely to develop technologies and secure sustainable competitiveness [18]. This might mean higher uncertainty, especially for companies with little knowledge or experience of using technology for innovation.

In this situation, acknowledging and understanding the current state of technology development, which has been conceptualized as technology sensing in the organization management literature [21], or termed as technology monitoring in the technology management literature [22,23], becomes an essential requirement for realizing the opportunities and evading the threats that the technology can bring about [4,19,23,24]. Accordingly, firms are increasingly interested in exploring external technologies, as well as managing internal technologies. A typical method for technology exploration is an expert-based approach, relying on the experts' knowledge and insights [25]. However, this approach has become not only time-consuming and expensive, but also invalid, as the technological range and topics to explore have become larger and have continued to change. Consequently, firms are in need of a more structured way to reduce the burden associated with the exploration of technological trends and gain quality information that companies need to support their innovation and strategic planning [26].

In response, much of the technology management literature has endeavored to develop methods and processes to facilitate technology trend exploration and monitoring. With the increasing amount of technology-related data, such as scientific publications and patent documents, significant developments have been made focusing on how to gain valuable information from such a large amount of data [27]. As scientific publications and patent documents are basically unstructured textual data, text mining has become a widely applied method to analyze textual data in fully- or semi-automated manners [28].

In particular, patents have been widely accepted as valuable data for the purpose of technology exploration with the following advantages. First of all, the patented technologies are worth analyzing because they can be assured of novelty and usefulness by the legal process [29]. In addition, the patented technologies are likely to have high technological and/or economic significance due to the patenting process being costly and taking years [30]. Also, patent documents contain bibliographic information such as filing and register dates, assignees, and prior art, thus enabling researchers to examine the characteristics of technologies from diverse perspectives like time, holders, and relationships among technologies. In addition, patent documents can be freely accessed [31] and online databases for patents are common these days. Hence, patent data can address the difficulty of obtaining objective data on technological development outside of the organization. Therefore, the appropriate analysis of patents data provides an organization with the capability to capture information regarding technology opportunities and threats [32].

Recent years have witnessed a significant increase in studies using patent data to capture technological information in other industrial areas, such as electronic shopping [33] and financial services [34]. In contrast, the utility of patent analysis for technology trend exploration has been relatively neglected in the logistics literature. From the perspective of technology adopters or technology users, research on important technologies and their main drivers and benefits of adoption in logistics has been discussed relying on experts' insights [5,9,20,35,36]. Many empirical studies have examined technology adoption [37,38] and its effect on logistics service performance [39]. However, few studies have taken methodological approaches to investigate logistics firms' technological competencies using patent data, except for a very limited number of studies, including Wu [40].

Under these considerations, this study aims to propose a topic modeling-based approach to patent analysis to explore technological trends in logistics. Latent Dirichlet allocation (LDA) is employed as a main method to extract hidden topics behind patent documents instead of analyzing metadata, such as

patent classification codes. The topics identified by LDA are further investigated regarding trends in patenting activity and major assignees for each topic. This helps to understand the technological landscape in logistics, both at the technological field-level and firm-level. The suggested approach is expected to offer implications of the use of patents for the purpose of exploring technology trends in logistics.

This paper is organized as follows. Section 2 presents a literature review of technology-driven innovation in logistics, patent analysis for technology exploration, and LDA for topic analysis. The suggested research methodology is explained in Section 3, and the analysis and results are presented in Section 4. Finally, the discussion and conclusions are provided in Section 5.

2. Background

2.1. Technology-Driven Innovation in Logistics

Like other industries, in logistics, innovation is imperative for firms to survive in the current knowledge-based economy and technology is a crucial requirement for innovation [16,41]. Technology enables not only the enhancement of the firms' core competency [42], but also the development and implementation of a new business model [43–45].

The emergence of new technologies is a crucial factor that has changed the scope of logistics [1]. Logistics these days usually entails purchasing, distribution, inventory management, packaging manufacturing, and customer services [46]. Above all, the Internet and information and communication technologies (ICTs) are incontrovertible enablers for logistics innovation by sharing information among all parties in a supply chain and connecting them [20]. Those technologies facilitate information flows and communication among partners, enabling global logistics processes [47]. On the other hand, the advent of the Internet has led to the growth of e-commerce and the emergence of diverse business models. As a consequence, such new business operations require different logistics activities [10] and strategies [48]. For instance, with the growth of e-commerce, global logistics firms such as FedEx and United Parcel Service (UPS) could become full-service logistics by securing the technological capability to coordinate the flow of goods and information within supply chains [19]. Online shopping, now common, opened new business opportunities for logistics firms that manage order fulfillment and distribution [15]. The recent development of mobile technologies and their applications has accelerated such trends because it has helped the transition from multi-channel to omni-channel environments. Where customers have multiple options to interact with firms and need a seamlessly integrated experience [49], logistics operations become complex and such roles are increasingly important to customer satisfaction and company performance [17].

While specific purposes and advantages differ for technologies, one of the fundamental values that technology-driven innovation can bring in logistics is the enhancement of the traceability of physical and information flows [50] and visibility across the whole supply chain [15,51]. Based on these benefits, the uptake of those technologies in logistics is usually sought for operational efficiency enhancement and cost reduction [52,53], ultimately aiming to optimize the whole supply chain [51]. Also, enhanced traceability and visibility can result in reducing the risks of uncertainty involved in the supply chain [54]. Much literature has shown the role of technologies in logistics in terms of operational efficiency [7,8] and improving customer satisfaction [1]. Some recent studies have called attention to the positive effect of technological capability on financial performance in logistics [55–57].

The importance of technology-driven innovation in logistics and SCM is easily found in practice. For example, Walmart has been well-known for its high investment in information technology related to real-time data collection, data warehousing, and computerized data exchange with upstream and downstream partners [58]. As a result, it has improved stock turnover [58], achieving cost competitiveness and a quick response [59]. A case study [60] of CJ-GLS, a Korean third-party logistics (3PL) provider, showed that the successful application of advanced technologies like RFID, aligned with a corporate strategy, could be a critical source of differentiation and competitiveness. Amazon,

which has become a 3PL provider by offering logistics services [61], goes a step further. In addition to efficiency enhancement and cost reduction in logistics operations, such as stocking, documentation, and picking/inspecting, Amazon has tried to predict consumers' demand at an individual level and implement anticipatory shipping by analyzing big data gathered from its well-connected supply chains [13].

Despite its apparent importance, it is only relatively recently that innovation has become a primary topic in logistics research [43,62]. Chapman et al. [16] reviewed the literature and theories related to service innovation in the logistics' context, and highlighted the adoption and creative deployment of new technology for logistics firms to innovate business models and gain competitive advantages. The strategic importance of technologies and innovation in logistics is also shown by the competition between FedEx and UPS over several decades. These firms have continuously adopted advanced technologies and innovation to enhance customer service and differentiate their competitive edge [63]. However, technological innovation remains overlooked in the logistics literature. For instance, Stock's [64] work of analyzing abstracts of 317 dissertations in logistics showed that few studies had dealt with technological innovation. Such little attention to technological innovation might partially be attributed to the practice that research and development (R&D) has less often been carried out by logistics service providers; technology has been adopted rather than developed [62]. Accordingly, most of the logistics literature has devoted a substantial effort to investigating the adoption and application of technologies in logistics, particularity focusing on ICTs. The main research topic of interest has been the effect of technologies on logistics performance [52]. Another main research theme is technological acceptance [37,38] and its effect on performance [39].

These previous studies have consistently emphasized the importance and effects of technologies in logistics and the need to prepare for changing technological advances [19,35]. In other words, the technology has become a strategic concern in logistics. Yet there is little research on how to become aware of technological options in logistics, except for relying on domain experts, which is a prerequisite for preparing for advances in an ever-growing technological environment.

Some researchers have tried to examine the technological trends in logistics by case studies, interviews, and surveys by domain experts. For example, Hammant [5] investigated information technology (IT) trends affecting the use of technology for supporting logistics and grouped the trends into four themes. The themes include integration and flexibility, electronic data interchange, hardware, and communication technology. Ellram et al. [35] investigated the current and future impact of IT on logistics in the retail industry, while focusing on electronic data interchange, point of sale systems, and barcoding, which were cutting-edge technologies at that time. Their findings were based on the survey of retail executives. Auramo et al. [20] identified, categorized, and prioritized key R&D topics in e-business logistics based on the interviews and the workshop with e-business and logistics experts. As a result, three categories of R&D topics were suggested, including information flow, supply network, and physical material flow, and four future research themes were proposed by combining different R&D topics. Harris et al. [9] discovered innovation opportunities and foresaw new business models with regard to future technological developments in freight multimodal transportation based on the analysis of EU framework programme projects. Using the Delphi research technique, Kache and Seuring [36] identified 43 opportunities and challenges related to big data analytics in supply chains. All these findings in the literature provide useful information for understanding the effect of technologies in logistics and in prospecting technological trends. However, the expert-based process can generate results that are biased toward the individuals' knowledge and experience. This risk can be more severe in the current circumstance in which technologies are rapidly developed and advanced technologies from different fields transform ways of competition. Also, the expert-based process for exploring technological trends is becoming costly as the number and areas of technologies to consider increase.

To sum up, the literature review on technology innovation in logistics has led to the conclusion that there has been a lack of research with the purpose of gaining insights into the whole technological

landscape in the logistics sector with the help of appropriate methods and objective data, concerning the following issues: what technological topics have evolved in logistics, how important they are, and who has competitive advantages and where. These are the main research questions of this paper.

2.2. Patent Analysis for Technology Exploration

A variety of items in a patent document can be grouped into structured data and unstructured data, depending on semantics and format [65]. The analytic methods differ depending on types of data [65]. On the one hand, structured data include the items that are uniform in both semantics and formats. The representative examples are patent number, filing date, register date, assignees, patent classification codes, and citations. The early research usually focused on structured data for decades [66,67]. A popular method for analyzing these data has been bibliometric analysis [68]. The specific methods are different depending on the subject of technology exploration. Some representative examples are: statistical analysis of bibliometric information to assess technological developments and competitiveness [69], co-assignee analysis to understand technological cooperation [70], co-classification analysis to reveal technological convergence patterns [71,72], and citation analysis to identify important patents or technology areas [73].

On the other hand, free texts of various lengths and contexts are classified as unstructured data, which include the title, abstracts, claims, and descriptions of the invention in patent documents. As the methods and algorithms to deal with textual data, such as text mining and natural language processing, have been developed, a lot of research on technological innovation has made use of textual data in patent documents [65]. While there is no single way of using patent texts for the purpose of technology exploration, a common task is to identify technology topics from the text data and investigate their characteristics. The specific methods for identifying technology topics vary from the keyword-based clustering of similar patent documents [74,75] to the topic modeling methods such as LDA [34,76–78]. By combining other methods like bibliometric analysis [30,70] and network analysis [75], more fruitful insights can be obtained.

To sum up, with the advantages of patents as a data source for innovation study and technology management and the advanced computerized methods like text mining, patent analysis has been a norm for technology exploration. However, few previous studies have performed patent analysis to assess technological innovation in logistics, except for Wu [40]. This study evaluated the performance of 3PL providers regarding technological innovation using patent statistics. The work showed that the patent statistics can provide a good picture of the state of logistics innovation and that logistics service providers can innovate by utilizing patent data appropriately and developing their technological strength. However, the study only focused on patent counts and citation counts according to the year among a range of information derivable from patents. It tried to identify and understand technological hot topics in logistics using International Patent Classification at the subclass level, but the pre-determined patent classification scheme has limitations in providing specific information about topics of patented technologies [79,80]. In contrast, the textual data in patent documents can provide richer information [81]. In addition, although analyzing the patents held by some 3PL providers can aid in assessing those firms' technological strength and performance, it is deficient for investigating logistics-related technologies thoroughly. It overlooked many technologies related to logistics not patented by logistics service providers. Accordingly, the subject of this study is the logistics-related patent, not the logistics service providers' patent, focusing on the textual data.

2.3. LDA for Topic Analysis

Given the growing amount of text data and the advanced research in machine learning, topic model-based approaches to automatically discover latent topics in documents have gained popularity. Specific algorithms differ from each other, but topic models commonly posit that a topic is assumed to contain a group of words that frequently occur together. In other words, text with different topics might use different vocabulary. Based on this assumption, topic models have been devised as probabilistic

models that attempt to find a set of words from the collection of documents that best explain the topics in documents. By discovering hidden relationships between words and topics across the collection of documents, topic models help us to identify documents that are similar in the use of words (i.e., topics) and annotate documents according to these topics [82]. Hence, the topic-based approach has provided a convenient way to organize, understand, and summarize a large collection of textual information in various fields [83].

Among several methods of topic modeling, LDA [83,84] is a Bayesian topic model and it has recently been widely used. Basically, LDA is a probabilistic model that assumes writing as the topic-word choice process [83,85]. Based on this assumption, LDA represents the documents as a mixture model with *k*-latent topics and each topic as multinomial distributions over words in documents [83,84]. The problem in identifying topics of documents is that we can only observe documents (i.e., the words contained in those documents), but the topic structure is hidden, such as topics per se, topic distributions in a document, word distributions for every topic, and topic assignment for every word in every document [83]. LDA resolves this problem by backtracking from the observed distribution of words in given documents to infer the hidden topic structure that is likely to have generated the collection of documents [83]. From a computational perspective, the inference means calculating the conditional distribution of hidden variables given the documents, called posterior distribution in Bayesian language.

Unlike other classical methods like the mixture of unigram models [86], LDA allows for a more realistic topic model by employing a mixed-membership model where the documents are simultaneously assigned to several topics to different degrees and the topic distributions vary over documents [87]. Another advantage of LDA stems from the fact that it is a complete probabilistic generative model [84,85]. Probabilistic Latent Semantic Indexing (pLSI), which was proposed earlier than LDA by Hofmann [88], shares the basic idea regarding the writing process and also represents a particular document as the mixing proportions for topics. However, pLSI suffers from providing no generative probabilistic model at the document level [84]. It is unclear how pLSI assigns probabilities to a document unseen in a training set, so it is difficult to generalize the model to new documents. Moreover, the number of parameters that pLSI needs to estimate increases linearly with the size of the collection of documents, causing a severe overfitting problem [84]. LDA overcomes such problems by introducing hidden random variables for the topic mixture weights for documents [84]. Namely, LDA is a complete generative model and easily deals with new documents. As a result, LDA is also free from the overfitting problem, as well as the increase of parameter numbers with the size of data.

Acknowledging the utility of LDA, recent studies used this method to discover technological topics and trends from scientific documents. Some examples are as follows. Lee et al. [78] predicted emerging topics given technological convergence by analyzing patents using LDA and link prediction techniques. Chen et al. [76] employed LDA to identify topics in patent claims and reveal their changes over time. Lee and Sohn [34] performed LDA on financial BM patents together with a survival analysis technique and then identified topic emerging patterns.

3. Research Methodology

As shown in Figure 1, our proposed approach consists of four steps: data collection, data pre-processing, topic identification, and topic exploration. The first and second steps are to collect patent documents related to logistics and then to prepare the collected documents in a form that can be analyzed by a topic-modeling algorithm. The third step is to identify the main topics of logistics-related patents. Finally, in the fourth step, the identified topics are further explored regarding trends in patenting activity and major assignees in each topic.
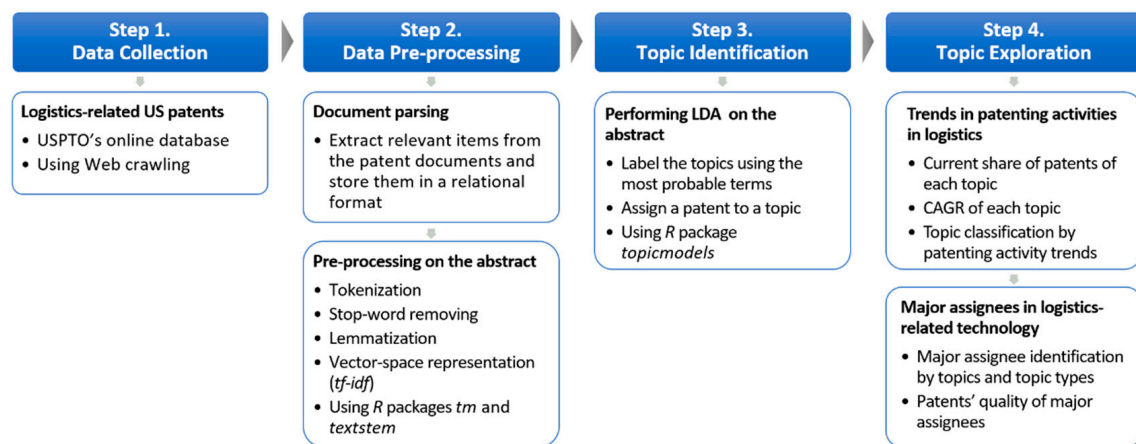
Figure 1. Overall process of the proposed approach.

3.1. Step 1: Data Collection

The patent documents related to logistics are collected from the United States Patent and Trademark Office (USPTO)'s online database. The data source is appropriate for exploring technological trends in logistics because it is a representative patent database [66] containing an enormous number of patents from all over the world and covers the most advanced technologies [89].

In previous studies on the industry sector-level patent analysis, patent classifications have often been used to find related patents (e.g., [34]). However, matching patent classifications to industry sectors poses a challenge [90]. Furthermore, a linkage between patent classifications and logistics remains unrevealed due to the lack of previous logistics research using patent data. Hence, a patent classification-based approach to finding target patents of this study risks missing important data.

Another option to determine target data is to find logistics service providers' patents. However, traditionally, logistics service providers have typically been technology adopters rather than technology developers [62]. Although the patenting activity by logistics service providers has grown since the 1990s, many of them still have small patent portfolios [40]. Instead, recent years have witnessed that firms in other fields, even those not considered related to logistics, are increasingly investigating and devoting efforts to securing logistics-related technologies recognizing the importance of logistics as a differentiator [18]. Given this situation, the firm-based search in which the patents held by logistics service providers are collected is not able to reflect such technological and industrial convergence trends in logistics.

The above-mentioned limitations of the classification- and firm-based approach to determine target patents can hinder achieving our purpose of investigating topical trends in logistic-related technologies behind patent data. Thus, this study employs a keyword-based approach: the logistics-related patents are defined as the patents that have the keywords *logistics* and *supply chain* in their title or abstract. There has been some confusion and debate on the definition and the usage of logistics and supply chain management [91]. A consensus on the relationship between logistics and supply chain management in both academia and practice is that supply chain management is more strategic and broader than logistics [92,93]. However, it is indisputable that both logistics and supply chain management concern the flow of products, services, and information [91]. In addition, when a patent application is written, the terms *logistics* and *supply chain* are not likely to be used in the strict sense as they are not technical terms. Hence, using supply chain as well as logistics as search keywords for logistics-related patents helps to reflect the scope of logistics that has been widened and blurred.

3.2. Step 2: Data Pre-Processing

Because the collected patent documents are in an unstructured text format, they should be pre-processed and transformed into a structured format for further analyses. The pre-processing procedure performed in this study is described in Figure 2.

Firstly, using the document parsing techniques, the relevant items are extracted from documents, such as the title, abstract, assignees, filing year, register year, classification code, and citation. Then, the items are stored in a relational format (i.e., a database table).

Among those items, the abstract is used as the input to LDA to identify topics because it typically includes the main problem that the patented technology addresses and the core of the technological solution. For this purpose, the abstract in a free-text format requires more pre-processing tasks with natural language processing techniques, including tokenization, lemmatization, stop-word removing, and vector-space representation.

First, through the tokenization process, every sentence of the abstract is broken into words, called tokens [94]. The resulting tokens go through the lemmatization process that reduces variants of a word to their 'lemma'. By doing this, a set of inflected forms of a word can be considered as a single item in the further analyses [95]. Then, the stop words that rarely contribute to representing the semantics of the documents are removed from the lemmatized words. There is no single universal list of stop words, but function words like articles, conjunctions, and prepositions are generally regarded as stop words. In addition, the common words of the language are also included in the stop-word list because they usually do not give any values to distinguishing documents [94,95]. Finally, the patent abstract is represented as a term frequency-inverse document frequency (TF-IDF) vector. The TF-IDF weighting scheme has been widely used to measure the importance of a given term in a document [96,97]. The TF-IDF value, an importance weight of term j in patent abstract *i*, is calculated by multiplying TF and IDF values. The formula is: $w_{i,j} = tf_{i,j} \times log(N/df_j)$, where $tf_{i,j}$ is the occurring frequency of term *j* in the patent abstract *i*; *N* is the total number of patent abstracts; and df_j is the number of patent abstracts that contain the term *j*. As this formula implies, the TF-IDF assumes that if a term is used a lot in a document and rarely appears in other documents, the term will likely have the ability to characterize the document, as well as to distinguish it from other documents. By incorporating the TF-IDF vector for every patent abstract, the TF-IDF matrix is generated, which is the input of the LDA in the next step.

3.3. Step 3: Topic Identification

The research question that this study aims to address by this step is: what technological areas have been developed in logistics? To this end, this study performs LDA to identify *k* major topics in logistics-related patents. Specifically, the LDA algorithm is applied to the TF-IDF matrix for the patent abstracts.

The LDA on the patent abstract is expected to generate two outputs: the probability that each patent document is related to each topic and the term distribution over each topic. Based on the per-topic distribution, each patent document is assigned to one of *k* topics exhibiting the highest probability. In addition to grouping patents with similar topic probability distributions, LDA also has the benefit of increasing the understanding of the latent topic structure by producing a term distribution over each topic. In this study, the top *n* words that are most likely to occur in each topic are used to label the *k* identified topics in the logistics-related patents.

[Relational table for collected patent documents]

patent no.	title	abstract	application no.	filling date	register no.	register date	assignee	assingee nation
US9139279	Systems and methods for long endurance airship operations	In one example, a long endurance airship system includes a payload airship and a first logistics airship mechanically joined to the payload airship to form a first combined airship, the payload airship and the logistics airship having design capabilities differing by ...	US13/623757	2012.09.20	US9139279	2015.09.22	STRATOSPHERIC AIRSHIPS, LLC	US
US8275803	System and method for providing answers to questions	A system, method and computer program product for providing answers to questions based on any corpus of data. The method facilitates generating a number of candidate passages from the corpus that answer an input query, an...	US12/152411	2008.05.14	US8275803	2012.09.25	International Business Machines Corporation	US
US8105413	Manufacturing of bioorganic-augmented high nitrogen-containing	The invention describes a new method for the production of high nitrogen organically-augmented inorganic fertilizer that incorporates municipal biosolids or organic sludges that can compete with traditional fertilizers such as ammonium phosphate, ammonium sulfate	US11/359751	2006.02.23	US8105413	2012.01.31	Vitag Corporation	US

1. Document parsing

2. Tokenization

Chopping a sentence up into pieces, called tokens

Input: Friends, Romans, Countrymen, lend me your ears;

Output: | Friends | Romans | Countrymen | lend | me | your | ears |

3. Lemmatization

better → good
walking, walked, walks → walk
meeting → meeting or meet
(depending on the context)

4. Stop-word removing

Dropping out some extremely common words which would appear to be of little value in helping select documents matching a user need are excluded.

patent no.	abstract
US913927	collecton of lemmatized tokens without stop-word
US827580	collecton of lemmatized tokens without stop-word
US810541	collecton of lemmatized tokens without stop-word

5. Vector-space representation

	abstract of patent	abstract of patent	abstract of patent	abstract of patent	abstract of patent
term 1					
term 2					
term 3					
term 4					
term 5					
...					

1) Term frequency (tf)
2) Inversed document frequency (IDF)

$$w_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right)$$

$tf_{i,j}$ = number of occurrences of i in j
df_i = number of documents containing i
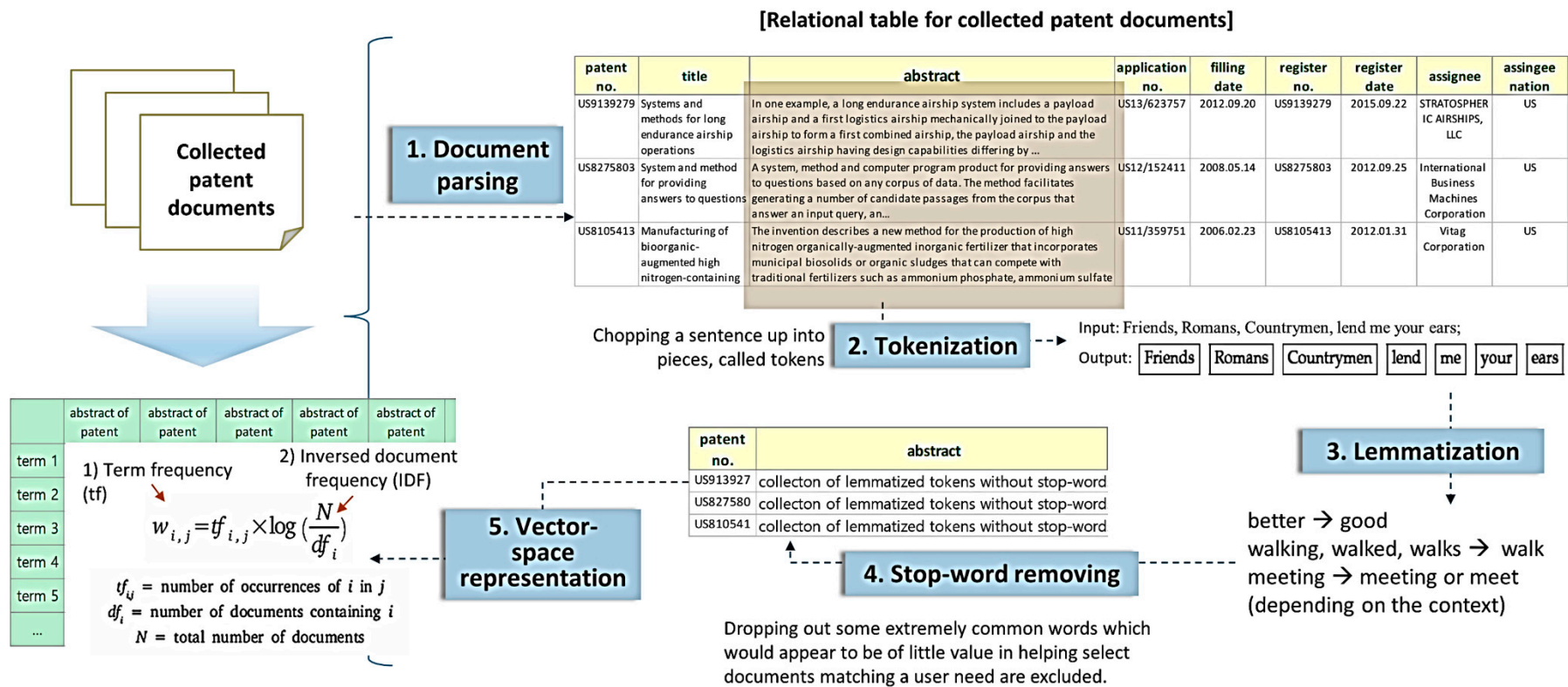N = total number of documents

Figure 2. Pre-processing procedure.

3.4. Step 4: Topic Exploration

In the last step, the identified topics are further explored from two aspects: trends in patenting activities over time and assignees in each topic. The research questions are as follows: how have technological areas changed over time? and who have been technological leaders in logistics? The investigation of these questions can offer the technological landscape in logistics at both the technological field-level and firm-level.

To this end, each topic is analyzed by synthesizing the information extracted from patents assigned to the topic. Among a variety of information that patents offer, the number of patent applications is the most basic for measuring the patent activity of a firm or of a technological field. Many other patent indexes for the attractiveness of technological fields and the technological ability of firms are based on it (for more information, see [98]). In this study, the number of patent application-based indexes are used for both topic-level and firm-level analyses.

On the one hand, to capture the patenting activity trends at the topic level, the current share of each topic in all logistics-related patents and the change in the patent share over time are examined. Specifically, the current patent share of a particular topic is measured as the number of patent applications in that topic divided by the total number of patent applications. Meanwhile, the change in the topic's patent share over time is quantified using the compound annual growth rate (CAGR). Using these two indexes as criteria, a classification framework for the identified topics is proposed in terms of patenting activity trends. These criteria capture two different aspects of patenting activity trends, i.e., current status and the growth rate. Thus, the proposed framework identifies four different types of technological topics. The *dominant topic* with a large patent share and positive CAGR of patent share, the *emerging topic* with a small patent share and negative CAGR of patent share, the *saturated topic* with a large patent share and negative CAGR of patent share, and the *declining topic* with a small patent share and the positive CAGR of patent share.

On the other hand, to identify technological leaders for each topic in logistics who have actively patented the related technologies, the relative technology share of a firm is used. A firm's technology share in a particular topic is the number of patents in the topic the firm has divided by the total number of patents in the topic [98]. Relative technology share is obtained by normalizing technology share to the value between 1 and 0, and thus it enables us to assess a firm's competitive position in the technological topic compared to leading firms [98]. Specifically, assignees with a relative technology share larger than 0.5 are screened out as major assignees in each topic.

It is reasonable that technology leaders are expected not only to perform much patent activity, but also to possess high-quality patents [99]. Also, many previous empirical studies found that there was more significant evidence for the positive relationship between the number of patent applications and the firms' performance when the patents' quality is considered together with the patent count [100–104]. Accordingly, to make the results of our analysis more meaningful, the patent quality indexes are also examined for major assignees in each topic. Four patent quality indexes are employed; the ratio of granted patents over a firm's total patent applications [100], technological scope calculated by the number of patent classifications [102], international scope defined as the size of the patent family [101], and forward citation frequency [103–105]. The former two indexes represent the technological quality of the firm's patent applications, while the latter two indexes represent the economic quality of the firm's patent applications [98].

4. Analysis and Results

4.1. Data and Pre-Processing

Employing the keyword-based approach to determine the target date, we found 3732 patent applications and collected them in HTML format by Web crawling for further analyses (search date of 4 April 2017). These patent documents were parsed to extract the relevant items such as the title, abstract, assignees, filing year, register year, classification code, citation, and the like. After being

formatted and reprocessed, these patent items were stored in a relational format, constructing the logistics-related patent database.

Except for the data having no abstract in the database, 3724 patents' abstracts were pre-processed by tokenization, lemmatization, stop-word removing, and vector-space representation. These pre-processing tasks on the abstracts were performed with an *R*-script. Specifically, tokenization and stop-word removing were performed using the *tm* package of *R* and lemmatization was implemented using the *textstem* package of *R*. The pre-processed patent abstracts were then transformed into a TF-IDF matrix. The specific tasks employed to construct the TF-IDF matrix were conducted using the *tm* package of *R*. We screened out the terms that did not appear in more than 95% of abstracts to exclude terms that rarely appeared. With the remaining 238 terms, consequently, the 3724 × 238 TF-IDF matrix was obtained.

Basic Features of Logistics-Related Patents

Before the results of topic identification by LDA are discussed, some basic features of our data are described, facilitating gaining insights into the technological development in logistics in terms of overall patenting activity. Firstly, as depicted in Figure 3, the number of logistics-related patent applications has drastically increased since 2000. It implies that the technology development related to logistics began to be concentrated in practice in the 2000s. This trend in patent applications might be in line with the consensus among previous studies about the significant influence of the ICTs on the change and innovation in logistics [106]. On the other hand, the steep drop in patent applications in 2016 is due to the patent filling process at the USPTO, which involves a significant time lag between the filing and disclosure or registration of applications, ranging from several months to over two years.
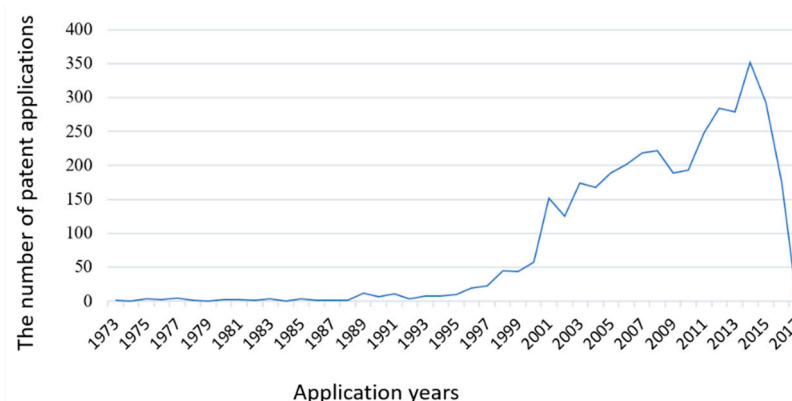

Figure 3. Distribution of the patent applications over application years.

Secondly, the assignee information allows us to know which major companies and institutions are involved in logistics patenting activities. As a result of the consolidation process that cleaned out misspellings and variations of the assignees' names, 1275 different assignees were identified. Some top assignees on the patenting activity are presented in Table 1. It is noteworthy that not only logistics companies, but also companies from various industries, filed logistics-related patents. In particular, IT-related companies are noticeable, such as International Business Machines, Microsoft, and Google. Electronic companies like Samsung, as well as Industrial equipment companies like Caterpillar, are also included. The representative retailing companies like Walmart, Alibaba, and Amazon in the top assignee list of our data are consistent with the current trends in retailing and logistics [18]. This result proved that it was reasonable not to choose a firm-based approach to collect the target patents.

Table 1. Top assignees of the logistics-related patents.

Rank	Assignee	# of Patent Applications	Cumulative %
1	INTERNATIONAL BUSINESS MACHINES CORP	130	3.5%
2	SAP SE	75	5.5%
3	MICROSOFT CORP	68	7.3%
4	GOOGLE INC	44	8.5%
5	UNITED PARCEL SERVICE OF AMERICA INC	41	9.6%
6	FEDEX CORP SERVICES INC	32	10.5%
7	GENERAL ELECTRIC CO	29	11.2%
8	XEROX CORP	27	12.0%
9	ACCENTURE GLOBAL SERVICES GMBH	25	12.6%
10	DEUTSCHE POST A G	24	13.3%
11	YAHOO INC	22	13.9%
12	BOEING CO	21	14.4%
12	SIEMENS AG	21	15.0%
14	AJINOMOTO CO INC	17	15.4%
14	LOCKHEED MARTIN CORP	17	15.9%
14	WAL MART STORES INC	17	16.3%
17	AT&T INTELLECTUAL PROPERTY I LP	16	16.8%
18	UNISYS CORP	15	17.2%
18	WABASH NATIONAL CORP	15	17.6%
20	ALIBABA GROUP HOLDING LTD	14	18.0%
20	INDUSTRIAL TECHNOLOGY RESEARCH INSTITUTE TAIWAN	14	18.3%
22	KONINKLIJKE PHILIPS NV	13	18.7%
22	LINKEDIN CORP	13	19.0%
22	UNITED STATES NAVY	13	19.4%
25	CATERPILLAR INC	12	19.7%
25	MICROSOFT TECHNOLOGY LICENSING LLC	12	20.0%
27	CERNER INNOVATION INC	11	20.3%
27	ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE	11	20.6%
27	STANFORD UNIVERSITY	11	20.9%
30	AMAZON TECHNOLOGIES INC	10	21.2%
30	ARROWSTREAM INC	10	21.4%
30	HONEYWELL INTERNATIONAL INC	10	21.7%
30	SAMSUNG ELECTRONICS CO LTD	10	22.0%

The United States Patent Classification Systems (USPC) categorizes every US patent document based on the technology claimed in the patent firstly into the class level and secondly into the subclass level (United States Patent and Trademark Office, December 2012). While a patent document can contain multiple classes, it mandatorily has one and only one primary classification at the subclass level which represents its core invention [107]. Even though only the primary classification is considered, our data reveal that various classes are related to the logistics-related patents in this study. At the class level, 187 different categories are identified from our data. The most frequent USPC class is the 705 (data processing: financial, business practice, management, or cost/price determination), and the class accounts for about 25% of our data. Namely, the other 75% of logistics-related patents cannot be captured by the 705 class. In terms of the USPC subclass, the logistics-related patents are more dispersed. A total of 1126 different USPC subclasses exist, and none of them account for more than 2% of our data. It implies that there is no dominant USPC subclass matching with the logistics-related patents. In addition to the subclasses having a name obviously related to logistics, such as 705/330 (shipping), 705/333 (tracking), and 235/385 (inventory), some subclasses that do not seem to be related to logistics at a glance, like 707/737 (clustering and grouping) and 235/375 (systems controlled by data bearing records), are also identified. In this situation, as mentioned in Section 3.1, the classification-based policy used to collect logistics-related patents is not suitable.

4.2. Technological Topics in Logistics

Based on the TF-IDF matrix for patent abstracts, this study performed LDA to identify major topics in logistics-related patents using the *topicmodels* package of *R*. There is no consensus on the parameter setting. In this study, after the experiments on different parameters, the number of topics was set at

10 because it provided substantial information, as well as ease of interpretation. Finally, the 10-topic model was fitted using Gibbs sampling to estimate the posterior distribution of latent topic structures. The *topicmodels* package of *R* offers the interface to implement an algorithm using Gibbs sampling suggested by Phan et al. [108]. As a result, each patent document was assigned to one of 10 topics that showed the highest probability. From the term distribution over topics, the top 10 words that are most likely to occur in each topic were selected. Finally, using these terms, the 10 identified topics in the logistics-related patents were labeled as presented in Table 2.

Table 2. Top assignees of the logistics-related patents.

Topic #	Most Likely Terms (Lemmatized)	Label
1	logistic, system, control, unit, communication, configure, vehicle, storage, operation, track	Vehicle communication
2	model, value, patient, risk, level, parameter, train, probability, predict, sample	Risk forecasting
3	order, service, item, customer, location, management, delivery, part, ship, inventory	Shipping service
4	present, determine, signal, individual, subject, measure, condition, detect, disclose, describe	Sensing
5	datum, associate, generate, andor, relate, event, analysis, source, compute, performance	Database system
6	process, component, module, form, material, cost, multiple, structure, support, member	Process optimization
7	less, network, comprise, store, supply, database, image, server, plan, chain	Supplier network
8	plurality, feature, select, identify, determine, input, score, result, function, group	Groupware
9	information, user, product, receive, application, interface, container, electronic, display, sensor	Container product information system
10	system, computer, time, embodiment, business, transaction, program, message, code, create	Business system

The topics include some conventional logistics areas like *shipping service* (topic 3) and *container product information system* (topic 5). *Supplier network* (topic 7) and *vehicle communication* (topic 1) are the areas that logistics has started considering more recently as its scope has widened. In addition, more general topics such as *risk forecasting* (topic 2), *sensing* (topic 4), *database system* (topic 5), *process optimization* (topic 6), *groupware* (topic 8), and *business system* (topic 10) are also identified. This might imply that general ICT development has been applied for the logistics purposes.

4.2.1. Trends in Patenting Activities in Logistics-Related Technology

The number of patent applications tends to sharply increase after the 2000s in all topics (see Figure 4a). This is predictable as such a tendency is the same as that of overall patent applications in logistics, presented in Figure 3. However, in terms of the relative share of patenting activity, the change patterns over time differ for different topics (Figure 4b). Hence, investigating the share of patents by topics and trends is more informative for understanding the recent technological focus on logistics than analyzing the mere number of patents.
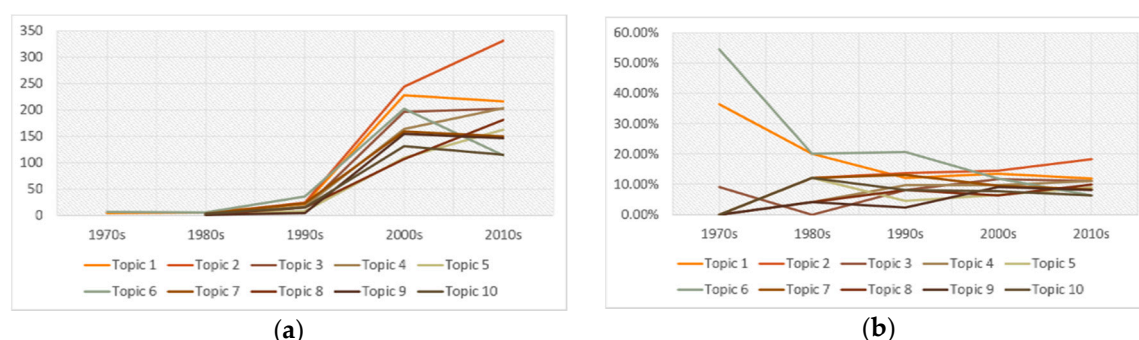
| (a) | (b) |

Figure 4. Trends in the number and the share of patents by topics. (**a**) Change in the number of patents over filing year; (**b**) Change in the share of patents over filing year.

There is a time lag between the filing and disclosure or registration of applications in USPTO's patent filing process. So, the number of patent applications in the last few years is usually excluded in the patent analysis. In this study, as Figure 3 depicts, the number of patent applications exhibited a decline after 2015. Thus, the base year for measuring the current patent share was chosen as 2014. Meanwhile, the time window for calculating the CAGR of the patent share was determined as the last 15 years between 2010 and 2014. This time period showed the drastically increasing number of patent applications in Figure 3, implying that patenting activities in logistics have become in earnest in the period. Table 3 summarizes the number of patent applications, the current patent share, and the CAGR of patent share by logistics-related topics.

Table 3. Share of patents by logistics-related technological topics.

Topic	The Number of Patent Applications	Current Patent Share [1]	CAGR of Patent Share [2]
Vehicle communication (T1)	474	12.76%	−1.27%
Risk forecasting (T2)	602	16.13%	−1.61%
Shipping service (T3)	413	12.03%	−1.19%
Sensing (T4)	384	10.07%	3.69%
Database system (T5)	282	6.84%	11.68%
Process optimization (T6)	363	10.05%	−6.07%
Supplier network (T7)	335	9.45%	2.52%
Groupware (T8)	303	7.09%	3.69%
Container product information system (T9)	305	8.19%	−0.21%
Business system (T10)	263	7.37%	−2.39%

[1] It is the current patent share in 2014. [2] It is the CAGR of patent share between 2010 and 2014.

Each topic was classified as one of four types of dominant, emerging, saturated, and declining topics, according to its current patent share and the CAGR of patent share. The average patent share in 2014 by topic was 10%, and it was used as a reference value to determine whether the current patent share of a particular topic is small or large. The reference value of the CAGR of patent share was set as 0. The result of classification is depicted in Figure 5.
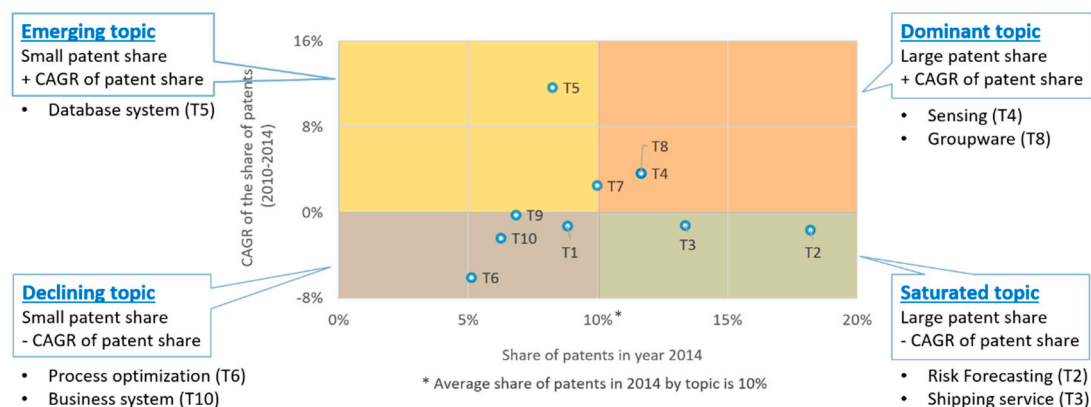
Figure 5. Classification of technological topics in logistics.

Sensing (T4) and *groupware* (T8) are classified as the dominant topics, having a current patent share that is larger than the average and a positive CAGR of patent share in the last 15 years. They can be considered the current main areas where technological development is actively achieved, and where future growth is also expected. *Database system* (T5) belongs to the emerging topic. This kind of technology area still has a relatively small patent share. However, since patenting activity has been increasing rapidly, it could be expected to grow in the future. *Risk forecasting* (T2) and *shipping service* (T3) are classified as saturated topics. The technology related to this kind of topic has been actively developed, but if trends continue, technological emphasis is likely to move onto the other topics. The last topic type includes declining topics that have a small patent share and negative CAGR of patent share. *Process optimization* (T6) and *business system* (T10) are included in this type, implying that these areas were the main subjects of innovation with technology in the past, but the interest in these areas has decreased in recent years. The remaining three topics—*vehicle communication* (T1), *supplier network* (T7), and *container product information system* (T9), are located at the border of the four types, and thus are not classified.

At first glance, it might seem awkward and counterintuitive that the *Database system* (T5) is classified as the emerging topic because the database is part of a well-established IT topic thread that started in the 1960s. However, the patents belonging to *Database system* (T5) are primarily focused on logistics because they are initially filtered out via searches using 'logistics' and 'supply chain' as keywords in step 1. As a result of this, it is possible to interpret that the use of technologies related to a database are receiving increased focus in logistics. This is in line with the practice of an increasing number of companies who are paying attention to the use of big data and a data-based approach to intelligent logistics [58,59,61]. In fact, there could be a lot of different sub-topics in databases within the logistics fields. The problem of sub-topics will be discussed in Section 5.3 on limitations.

4.2.2. Major Assignees in Logistics-Related Technology

For the 10 technological topics in logistics, major assignees having a relative technology share larger than 0.5 were identified and their average patent quality indexes were also calculated, as shown in Table 4.

The value of a relative technology share larger than 0.5 means the major assignees in each topic have more than 50% of the number of patents that the top assignee in each topic holds. In other words, major assignees identified in this study are assignees whose technological distance to the leading firm for the topic is less than 0.5. The number of major assignees for each topic is at least one, and at most five. In the case of the topic *groupware* (T5), only one major assignee (International Business Machines) was identified, meaning that it is less likely that there are strong competitors related to groupware technology for logistics. In contrast, a total of five major assignees, which is the greatest

value, were identified for the topic *shipping service* (T3). This implies that shipping service-related technologies have been competitively developed and patented by those major assignees.

Some transportation and warehousing companies showed remarkable patenting activities in traditional logistics areas. UPS and FedEx turned out to be major players regarding patenting activities in *shipping service* (T3). *Container product information system* (T9) is another area that logistics companies like FedEx have devoted their effort to in order to secure a technological advantage. Although Deutsche Post shows technological advantages in *supplier network* (T7) and UPS's efforts to technology development for *vehicle communication* (T1) is noticeable, for the reaming topics other than *shipping service* (T3) and *container product information system* (T9), many of the major technological leaders turned out not to be logistics companies, but to be IT-related companies.

As for the patent quality indexes, most of the major assignees show above-average values. This means that the major assignees have not only many patents, but also some level of quality patents. Interestingly, the logistics companies' patents tend to have much higher values of the international scope than the average. The international scope of a certain assignee was calculated by the average size of the patent family and the family size is the number of countries in which the patents are filed for the same invention [101]. The reason for an assignee having the patent family is that it has to file the patent in each country where it plans to perform business using the invention. Because filing and maintaining patents requires much expense, the family size is related to the expected value of innovation [101]. Considering the definition and characteristics of patent family size and the fact that UPS and FedEx are the representative global logistics companies, the high values of the international scope of logistics companies' patents imply that their patenting activities are likely to pursue business worldwide using the patent-protected invention.

We also organized the major assignees according to the four types of topics, as depicted in Figure 6. This enables us to examine a major assignee's portfolio of logistics-related patenting technology. For instance, International Business Machines (IBM) possess all types of logistics-related technologies, whereas Microsoft has developed logistics-related technologies in dominant, emerging, and saturated topics. Logistics companies including FedEx and UPS are only technology leaders in saturated topics. Although additional information is obviously needed to reach some conclusion, they seem to need to secure the balance of their technology portfolios to prepare for the future. SAP has logistics-related patents belonging to declining topics. More analyses and information are required, such as those regarding core competency and business strategy, to reach a clear conclusion, but our results suggest the possibility that a new portfolio strategy can be helpful for SAP to demonstrate technical dominance in the logistics-related business.

Dominant topics		Emerging topics	Saturated topics		Declining topics	
Sensing (T4)	Groupware (T8)	Database system (T5)	Risk forecasting (T2)	Shipping service (T3)	Process optimization (T6)	Business system (T10)
AJINOMOTO CO INC	INTERNATIONAL BUSINESS MACHINES CORP	INTERNATIONAL BUSINESS MACHINES CORP	MICROSOFT CORP	UNITED PARCEL SERVICE OF AMERICA INC	SAP SE	SAP SE
MICROSOFT CORP		MICROSOFT CORP	INTERNATIONAL BUSINESS MACHINES CORP	INTERNATIONAL BUSINESS MACHINES CORP	INTERNATIONAL BUSINESS MACHINES CORP	
			GOOGLE INC	FEDEX CORP SERVICES INC		
				GOOGLE INC		
				ACCENTURE GLOBAL SERVICES GMBH		

Figure 6. Major assignees by topic types.

Table 4. Top assignees of the logistics-related patents.

Topic	Major Assignees	Relative Technology Share	Ratio of Registered Patents	Technological Scope	International Scope	Forward Citation Frequency
Vehicle communication (T1)			0.57	2.19	0.51	6.91
	SAP SE	1.00	0.82	2.91	6.18	8.56
	UNITED PARCEL SERVICE OF AMERICA INC	0.82	0.78	2.56	13.44	5.29
	ARROWSTREAM INC	0.82	0.78	1.11	11.56	23.86
	GENERAL ELECTRIC CO	0.55	0.83	3.33	3.83	6.80
Risk forecasting (T2)			0.21	1.93	0.45	5.92
	MICROSOFT CORP	1.00	0.91	2.43	3.22	14.06
	INTERNATIONAL BUSINESS MACHINES CORP	0.78	0.44	2.33	1.94	7.00
	GOOGLE INC	0.57	0.77	2.85	3.38	10.00
Shipping service (T3)			0.48	1.90	0.34	5.84
	UNITED PARCEL SERVICE OF AMERICA INC	1.00	0.31	2.06	8.63	28.40
	INTERNATIONAL BUSINESS MACHINES CORP	0.94	0.47	2.27	2.53	25.43
	FEDEX CORP SERVICES INC	0.81	0.00	0.00	28.23	0.00
	GOOGLE INC	0.69	0.36	1.00	2.27	5.00
	ACCENTURE GLOBAL SERVICES GMBH	0.56	0.78	3.44	7.44	20.57
Sensing (T4)			0.45	2.01	0.46	7.58
	AJINOMOTO CO INC	1.00	0.31	1.56	5.81	2.00
	MICROSOFT CORP	0.63	0.70	2.80	2.70	4.00
Database system (T5)			0.48	1.79	0.45	6.53
	INTERNATIONAL BUSINESS MACHINES CORP	1.00	0.46	0.92	1.46	7.33
	MICROSOFT CORP	0.77	0.90	1.90	1.90	7.00
Process optimization (T6)			0.59	2.49	0.53	10.51
	SAP SE	1.00	0.81	2.62	1.33	18.75
	INTERNATIONAL BUSINESS MACHINES CORP	0.62	0.92	4.00	3.38	26.25
Supplier network (T7)			0.58	2.19	0.47	11.94
	INTERNATIONAL BUSINESS MACHINES CORP	1.00	0.67	3.08	2.67	45.43
	DEUTSCHE POST A G	0.50	0.00	1.00	3.00	0.00
	UNISYS CORP	0.50	0.00	0.00	1.00	0.00
Groupware (T8)			0.44	1.69	0.45	7.73
	INTERNATIONAL BUSINESS MACHINES CORP	1.00	0.56	1.78	4.81	21.44
Container product information system (T9)			0.41	1.87	0.36	6.52
	INTERNATIONAL BUSINESS MACHINES CORP	1.00	0.53	1.53	1.93	6.88
	FEDEX CORP SERVICES INC	0.87	0.31	0.62	9.00	10.75
	SAP SE	0.53	0.88	3.13	1.63	9.14
Business system (T10)			0.41	1.91	0.39	13.81
	SAP SE	1.00	0.68	1.79	2.00	22.75

5. Discussion and Conclusions

5.1. Discussion

5.1.1. Patent Search Policy for Investigating Technological Trends in Logistics

This study employed a keyword-based approach to finding patents' data to be analyzed, instead of classification-based and firm-based approaches. The analysis results showed that the assignees in the logistic-related patents were from different industrial fields and there was no dominant USPC class, thereby confirming the advantage of our approach to targeting patent data for investigating technological trends in logistics.

In addition, it is worth noting that this study did not limit the target data to the business method patents due to the characteristics of logistics. Previous studies analyzing patents to identify technological topics in the specific service fields have usually limited their target data to business method patents. For example, Lee et al. [33] and Lee and Sohn [34] analyzed the U.S. business method patents in the electronic shopping services and financial service, respectively. Business method patents include information on business strategies and associated technological configurations at an operational level [109], and thus can be a valuable source of information on how to use technologies as innovative means. In the USPC system, the U.S. business method patents correspond to class 705 (*data processing: financial, business practice, management, or cost/price determination*). In fact, the sub-categories of class 705 include logistics-related areas such as inventory management. However, as its name implies, the technology that a business method patent in class 705 can cover is limited to that related to the machines and corresponding methods for performing data processing or calculation operations [110]. This means that exploring business method patents might entail only providing information about software technologies, overlooking non-software technologies. However, logistics-related technologies are highly likely to include both software and non-software technologies because logistics entails both intangible and tangible elements according to the definition and the scope of logistics [1]. In addition, a new software technology that is still not used to generate new business methods at that time of analysis could not be captured by the analysis of business method patents, even if promising business methods can be created based on that technology. Thus, for a comprehensive analysis, the keyword-based approach to finding target patents was implemented for all utility patents in the USPTO database, rather than business method patents. The analysis result of this study showed that the most frequent USPC class in the logistics-related patents was the 705, but its share was only 20%. The other 75% of logistics-related patents could not be captured by class 705. This result also supports our policy.

5.1.2. Usefulness of a Tool for Technology Trend Exploration in Logistics

Although logistics service providers have typically been technology adopters rather than technology developers [62], understanding the technological landscape in logistics is still important in their strategies and competencies. For example, logistics service providers with a low technological ability, or in the business development stage, can leverage the cooperation with customer firms with high technology to facilitate knowledge and technology transfer between companies [1]. This kind of partnership can even enable logistics service providers to generate a new, technology-based business model [43]. A prerequisite for making this strategy succeed is to be aware of technological trends outside of the organization. In this situation, the suggested approach in this study can help logistics organizations to find candidate partnering firms by providing information about who has competitive advantages in logistics-related technological topics.

In addition, with the rapid development of technologies and the rapid change in the business environment as a consequence, logistics is no longer an area of merely technology application. Increasing startup companies in logistics have invested in compelling technologies as their core competency, and some of them have begun to attract global logistics companies' interest [111]. The big

companies outside the logistics industry, such as Amazon, have also received great attention for their remarkable development of logistics-related technologies [112]. It was also observed in this study that there are increasing patenting activities related to logistics. In this circumstance, to take advantage of opportunities and respond to the threats that technological shifts can cause, organizations need to proactively explore technological trends [21]. This means that the exploration of technological trends should be performed more thoroughly and regularly, requiring more resources and time. In this regard, the suggested approach in this study can effectively aid the extensive tasks associated with technological trends exploration in logistics because it is based on a computerized method and semi-automated processes to collect patent data, pre-process the data, identify technological topics, and examine technological trends per topic.

In the organization management literature, it has been recognized empirically that a large organization is favorable to technology sensing for acquiring knowledge about and understanding new technological development [113]. Thus, as a supporting tool for technological trend exploration, the suggested approach in this study would be helpful for small logistics companies that are the majority of the logistics industry. However, it does not mean the suggested approach is useless for large organizations. The positive relationship between firm size and technology sensing asserted in the literature is based on the premise that technology sensing relies on individuals' knowledge. In practice, however, few individuals get directly involved in technological-sensing tasks via the tasks-force or the dedicated department. Thus, even for large firms, the exploration of technological trends is challenging without supporting tools, because technology development advances rapidly and the development patterns are unprecedentedly complicated. Therefore, the proposed method would be useful regardless of company size.

5.1.3. Patent-Based Approach to the Exploration of Technology-Driven Innovation in Logistics

By definition, only novel and nonobvious technical invention can be patented [29]. Also, to be patented, the technical invention should be useful; meaning it can be practically applied for use in industries [29]. Therefore, the patent-based approach enables practitioners and researchers to investigate technological innovation with a high potential for industrial applicability. In addition, the availability of patent data [31] enables an organization to access external innovation [32]. Due to these features, patents are reliable data for this paper and are well-aligned with its primary focus: the development of an intelligent tool to support technology trend exploration in logistics.

When it comes to the specific types of technological innovation, a popular classification scheme is incremental and radical innovations. The radicalness of innovation is usually defined within a continuum with respect to the novelty and impact of innovation [114]. Due to the novelty, non-obviousness, and usefulness requirements for patents [29], the analysis of patents is more likely to provide information on innovation that presents radical steps forward or completely new approaches to problems, rather than incremental improvements that branch off and build on prior ideas. There are two different views on the patent-based measurement of how radical an innovation is. The first view assumes that individual patents within a specific group and timeframe have a similar level of impact and novelty. Accordingly, this approach assesses the radicalness of a group of patents at a certain time by employing the construct of an S-curve technology life cycle [115]. Specifically, patent counts in a group over time are fitted to an S-curve, which are then separated into introduction, growth, mature, and declined phases. The earlier a group of patents is plotted on this phase graph, the more radical the related technological innovation is considered to be [116].

The suggested approach in this paper is related to the first view. We did not fit patent counts to the s-curve directly, but the four topic types suggested in this study can correspond to the four phases of the technology life cycle. The two criteria, i.e., the current patent share and CAGR of patent share, imply the performance and the growth rate of the topic, respectively. Namely, the emerging topic with a small patent share and negative CAGR of patent share could correspond to the introduction phase; the dominant topic with a large patent share and positive CAGR of patent share could be matched

with the growth phase; the saturated topic with a large patent share and negative CAGR of patent share can be regarded as the mature phase; and the declining topic with a small patent share and the positive CAGR of patent share can be considered to enter the declined phase.

The second view posits that individual patents should have different levels of novelty and impact, even if they are applied at a similar time. The approach based on this view measures the radicalness of innovation at a patent level, mostly using citation-based indices and the content of invention [117]. Because this paper focuses on finding groups of similar patents according to their invention content (i.e., technological topics and the investigation of group-level characteristics of logistics-related technologies), identifying which patents are more radical within a topic is beyond the scope of this paper. However, the inclusion of the evaluation of individual patents could be helpful for some cases, like the benchmarking of technologies.

5.1.4. Time Window of Patent Data for Topic Identification and Exploration in Logistics

The topics behind logistics-related patents discovered by LDA could be different depending on the choice of time windows of subject patents. For example, if the time window is limited to the latest patents, only the patents related to the latest technologies in logistics are considered. Hence, if the focus is on capturing the cutting-edge technologies at the time of analysis, using the latest patents as base data for the topic identification might provide better results. In contrast, this study used the whole set of collected patents. However, the impact of the time window on the topic identification would not be significant for this study because the patents filed after 2000 account for around 94% of the whole set of patents.

In addition, the limited time window provides less information about the changes in technological topics over time. Because this study aims to reveal not only what technological topics are in the logistics fields, but also how related technological topics have been changed, using the latest patents is not appropriate. On the other hand, a wide time window can entail the risk of considering and assigning undue importance to obsolete or outdated topics. To mitigate such a risk, in this study, the patent data were narrowed down to the latest 15 years for topic exploration. Therefore, by employing different time windows for the topic identification and exploration, this study could provide information on historical changes of technological topics in logistics, as well as the current status of the topics.

5.2. Contributions

This study contributes to mitigating gaps concerning the lack of logistics literature on innovation [39,54] and use of patent data for the innovation study [36]. In particular, textual data of patent documents, which have rarely been used in logistics research despite their value, were analyzed. In addition, the logistics-related patents were defined based on the keywords rather on patent assignees or patent classifications. As such, this study could investigate technological trends in logistics more thoroughly and provide specific information about topics behind patents. This study also contributes to the logistics innovation research from the perspective of research methods by suggesting a topic-modeling approach to patent data with LDA, which is an advanced algorithm for analyzing textual data. Given the increase of logistics-related patents, the utility of our topic-modeling-based approach is expected to increase because it can handle a large amount of unstructured text data efficiently. In addition, LDA enables us to group and analyze patent documents that show similar word-use patterns, called topics. Although, in some aspects, the result seems similar to that of clustering algorithms, clustering algorithms are based on a black-box mechanism. In contrast, LDA provides a way of interpreting the patent documents grouped in a topic by explicitly representing the topics using word probability distributions over topics [75].

The results of our suggested approach can enhance the understanding of what technology areas are important in logistics, who the important players in each area are, and how they have changed over time. The 10 identified topics in logistics-related patents included not only areas traditionally regarded as a logistics area, such as *shipping service* and *container product information system*, but also

industry-independent areas, such as *database system*. This study made an attempt to examine dynamic changes of technological topics in logistics, as well as the current status. The classification scheme using the current patent share and the CAGR as criteria generated four different topic types: dominant, emerging, saturated, and declining. As a result, the classification scheme enables us to consider that different technologies have lead innovation in logistics in different ages when assessing a particular technological topic. As a result of this study, while traditional logistics areas were classified as saturated and declining topics, *database* and *sensing* were classified as emerging topics. This implies that logistics is also affected by rapid technological development in data acquisition and analysis. Also, it enables the anticipation of the importance of a specific technological area in the future. Further, it implies the need for different patenting strategies depending on changing patterns.

This study found that many major assignees of logistics-related patents were from various industries. For instance, UPS and FedEx are technological leaders among the logistics companies, but their technology competency tends to be limited to traditional logistics areas. More IT companies seem to be likely to expand their influence in logistics. This reflects the expansion and convergence trends in the logistics field and implies the increasing possibility of competition, as well as cooperation, from unexpected industries and technology sectors.

As the research on investigating logistics-related technologies and the resulting innovations using patent data is still in the infant stage and the logistics exhibits converging features, the keyword-based approach to securing the target patent data can be more useful than the firm- and classification-based ones. The USPC classes associated with the topics identified in this study can be a starting point for researchers trying to investigate the US patents' classification code related to logistics. Also, the major assignees by the logistics-related technological topics can help researchers and practitioners discover and monitor the technological leaders in particular areas of interest.

5.3. Limitations

This study has limitations which require future research. First, this study can be improved by adjusting data collection policy. This study took a keyword-based approach to define logistics-related patents, recognizing the limitations of classification-based and firm-based approaches. However, the use of keywords to find target patents, by its nature, suffers from the drawback that the quality and quantity of the data largely and inevitably depend on the search keywords and search queries. This study used the two keywords *logistics* and *supply chain* to determine logistics-related patents, but the resulting patents are not likely to cover all patents that are related to logistics. For instance, the data collected in this study only include 10 patents applied by Amazon. However, in recent years, a lot of media have reported on Amazon's increasing patent application activities in logistics-related areas (for example, see [104]). It indicates the high possibility that the keywords *logistics* or *supply chain* are not included in some patents, even if their patented inventions are actually related to logistics. Therefore, the data collection policy needs to be improved, and the data collected in this study could be utilized as a good starting point. For example, the words frequently occurring with *logistics* and *supply chain* in our data can be candidates in specifying search-keywords for logistics-related patents. Also, the classification codes and assignees derived from our data can pave the way for the keyword-based approach to be complemented by the classification-based or firm-based approach to finding logistics-related patents.

Second, this study can also be elaborated from the analysis aspect. The identified technological topics and the analysis of trends in patenting activities and assignees in each topic help in illuminating the technological landscape in logistics to some degree. However, there is still room for improvement by subdividing the identified topics. Each topic discovered in this study included around 200 to 600 patent documents. These might protect different technologies regarding specific purposes and utilization. Thus, it can provide richer information to break down the topics into detailed subtopics and to analyze the trends at the subtopic level. In addition, the subtopic-level analysis is more suitable to investigate the technological landscape according to the business model components or value chain

components in logistics, enabling the identification opportunities of technology-based innovation in logistics.

References

1. Gunasekaran, A.; Ngai, E.W.T. The successful management of a small logistics company. *Int. J. Phys. Distrib. Logist.* **2003**, *33*, 825–842. [CrossRef]
2. Sandberg, E. Understanding logistics-based competition in retail–a business model approach. *Int. J. Retail Distrib. Manag.* **2013**, *41*, 176–188. [CrossRef]
3. Tracey, M.; Limm, J.S.; Vonderembse, M.A. The impact of supply-chain management capabilities on business performance. *Supply Chain Manag.* **2005**, *10*, 179–191. [CrossRef]
4. Bowersox, D.J. Improving The Logistics Marketing/Sales Interface. In *Annual Conference Proceedings of the Council of Logistics Management*; Council of Logistics Management: Oak Brook, IL, USA, 1991; pp. 245–255.
5. Hammant, J. Information technology trends in logistics. *Logist. Inf. Manag.* **1995**, *8*, 32–37. [CrossRef]
6. Renko, S.; Ficko, D. New logistics technologies in improving customer value in retailing service. *J. Retail. Consum. Serv.* **2010**, *17*, 216–223. [CrossRef]
7. Bourlakis, M.; Bourlakis, C. Integrating logistics and information technology strategies for sustainable competitive advantage. *J. Enterp. Inf. Manag.* **2006**, *19*, 389–402. [CrossRef]
8. Brah, S.A.; Lim, H. The effects of technology and TQM on the performance of logistics companies. *Int. J. Phys. Distrib. Logist.* **2006**, *36*, 192–209. [CrossRef]
9. Harris, I.; Wang, Y.; Wang, H. ICT in multimodal transport and technological trends: Unleashing potential for the future. *Int. J. Prod. Econ.* **2015**, *159*, 88–103. [CrossRef]
10. Yu, Y.; Wang, X.; Zhong, R.Y.; Huang, G.Q. E-commerce logistics in supply chain management: Practice perspective. *Procedia Cirp* **2016**, *52*, 179–185. [CrossRef]
11. Zhu, X.; Mukhopadhyay, S.K.; Kurata, H. A review of RFID technology and its managerial applications in different industries. *J. Eng. Technol. Manag.* **2012**, *29*, 152–167. [CrossRef]
12. Fera, M.; Macchiaroli, R.; Fruggiero, F.; Lambiase, A.; Miranda, S. Application of a business process model (BPM) method for a warehouse RFId system implementation. *Int. J. Technol.* **2017**, *8*, 57–77. [CrossRef]
13. Agrahri, H.; Ahmed, F.; Verma, V.K.; Purohit, J.K. Benefits of implement big data driven supply chain management: An ISM based model. *Int. J. Eng. Sci.* **2017**, *7*, 11426–11431.
14. Fera, M.; Fruggiero, F.; Lambiase, A.; Macchiaroli, R.; Miranda, S. The role of uncertainty in supply chains under dynamic modeling. *Int. J. Ind. Eng. Comput.* **2017**, *8*, 119–140. [CrossRef]
15. Alshawi, S. Logistics in the Internet age: Towards a holistic information and processes picture. *Logist. Inf. Manag.* **2001**, *14*, 235–242. [CrossRef]
16. Chapman, R.L.; Soosay, C.; Kandampully, J. Innovation in logistic services and the new business model: A conceptual framework. *Manag. Serv. Qual. Int. J.* **2002**, *12*, 358–371. [CrossRef]
17. Drabik, P.; Zamecnik, P. Key Aspects of Logistics for Online Store and Multichannel Distribution. In Proceedings of the 16th International Joint Conference: Central and Eastern Europe in the Changing Business Environment, Prague, Czech Republic, 27 May 2016; pp. 97–111.
18. Hofmann, E.; Osterwalder, F. Third-party logistics providers in the digital age: Towards a new competitive arena? *Logistics* **2017**, *1*, 9. [CrossRef]
19. Olavarrieta, S.; Ellinger, A.E. Resource-based theory and strategic logistics research. *Int. J. Phys. Distrib. Logist.* **1997**, *27*, 559–587. [CrossRef]

20. Auramo, J.; Aminoff, A.; Punakivi, M. Research agenda for e-business logistics based on professional opinions. *Int. J. Phys. Distrib. Logist.* **2002**, *32*, 513–531. [CrossRef]

21. Srinivasan, R.; Lilien, G.L.; Rangaswamy, A. Technological opportunism and radical technology adoption: An application to e-business. *J. Mark.* **2002**, *66*, 47–60. [CrossRef]

22. Gerken, J.M.; Moehrle, M.G. A new instrument for technology monitoring: Novelty in patents measured by semantic patent analysis. *Scientometrics* **2012**, *91*, 645–670. [CrossRef]

23. Teichert, T.; Mittermayer, M.A. Text Mining for Technology Monitoring. In Proceedings of the International Engineering Management Conference, Cambridge, UK, 18–20 August 2002; Volume 2, pp. 596–601.

24. Basberg, B.L. Patent statistics and the measurement of technological change—An assessment of the Norwegian patent data, 1840–1980. *World Patent Inf.* **1984**, *6*, 158–164. [CrossRef]

25. Lichtenthaler, E. Technological change and the technology intelligence process: A case study. *J. Eng. Technol. Manag.* **2004**, *21*, 331–348. [CrossRef]

26. Porter, A.L.; Cunningham, S.W. *Tech Mining: Exploiting New Technologies for Competitive Advantage*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2005.

27. Kerr, C.I.; Mortara, L.; Phaal, R.; Probert, D.R. A conceptual model for technology intelligence. *Int. J. Technol. Intell. Plan.* **2006**, *2*, 73–93. [CrossRef]

28. Hashimi, H.; Hafez, A.; Mathkour, H. Selection criteria for text mining approaches. *Comput. Hum. Behav.* **2015**, *51*, 729–733. [CrossRef]

29. United States Patent and Trademark Office. General Information Concerning Patents. 2015. Available online: https://www.uspto.gov/patents-getting-started/general-information-concerning-patents (accessed on 7 March 2018).

30. Basberg, B.L. Patents and the measurement of technological change: A survey of the literature. *Res. Policy* **1987**, *16*, 131–141. [CrossRef]

31. Daim, T.U.; Rueda, G.; Martin, H.; Gerdsri, P. Forecasting emerging technologies: Use of bibliometrics and patent analysis. *Technol. Forecast. Soc.* **2006**, *73*, 981–1012. [CrossRef]

32. Porter, A.L.; Detampel, M.J. Technology opportunities analysis. *Technol. Forecast. Soc.* **1995**, *49*, 237–255. [CrossRef]

33. Lee, C.; Park, H.; Park, Y. Keeping abreast of technology-driven business model evolution: A dynamic patent analysis approach. *Technol. Anal. Strateg.* **2013**, *25*, 487–505. [CrossRef]

34. Lee, W.S.; Sohn, S.Y. Identifying Emerging Trends of Financial Business Method Patents. *Sustainability* **2017**, *9*, 1670.

35. Ellram, L.M.; La Londe, B.J.; Weber, M.M. Retail logistics. *Int. J. Phys. Distrib. Logist.* **1999**, *29*, 477–494. [CrossRef]

36. Kache, F.; Seuring, S. Challenges and opportunities of digital information at the intersection of big data analytics and supply chain management. *Int. J. Oper. Prod. Manag.* **2017**, *37*, 10–36. [CrossRef]

37. Lin, T.H.; Lin, I.C. Factors for information technology acceptance willingness and adoption in logistics industry from supply chain perspectives. *Int. J. Electron. Bus.* **2014**, *12*, 167–177.

38. Ramanathan, R.; Ramanathan, U.; Ko, L.W.L. Adoption of RFID technologies in UK logistics: Moderating roles of size, barcode experience and government support. *Expert Syst. Appl.* **2014**, *41*, 230–236. [CrossRef]

39. Evangelista, P.; Mogre, R.; Perego, A.; Raspagliesi, A.; Sweeney, E. A survey based analysis of IT adoption and 3PLs' performance. *Supply Chain. Manag.* **2012**, *17*, 172–186. [CrossRef]

40. Wu, Y.-C.J. Assessment of technological innovations in patenting for 3rd party logistics providers. *J. Enterp. Inf. Manag.* **2006**, *19*, 504–524.

41. Kandampully, J. Innovation as the core competency of a service organisation: The role of technology, knowledge and networks. *Eur. J. Innov. Manag.* **2002**, *5*, 18–26. [CrossRef]

42. Closs, D.J.; Goldsby, T.J.; Clinton, S.R. Information technology influences on world class logistics capability. *Int. J. Phys. Distrib. Logist.* **1997**, *27*, 4–17. [CrossRef]

43. Flint, D.J.; Larsson, E.; Gammelgaard, B.; Mentzer, J.T. Logistics innovation: A customer value-oriented social process. *J. Bus. Logist.* **2005**, *26*, 113–147. [CrossRef]

44. Kerr, A. Information technology–creating strategic opportunities for logistics. *Int. J. Phys. Distrib.* **1989**, *19*, 15–17. [CrossRef]

45. Stock, J.R. Managing computer, communication and information technology. *Logist. Transp. Rev.* **1990**, *26*, 133.

46. Bowersox, D.J.; Closs, D.J. *Logistical Management: The Integrate Supply Chain Management*; McGraw-Hill: London, UK, 1996; pp. 63–70.

47. Lewis, I.; Talalayevsky, A. Third-party logistics: Leveraging information technology. *J. Bus. Logist.* **2000**, *21*, 173–186.

48. Delfmann, W.; Albers, S.; Gehring, M. The impact of electronic commerce on logistics service providers. *Int. J. Phys. Distrib. Logist.* **2002**, *32*, 203–222. [CrossRef]

49. Lazaris, C.; Vrechopoulos, A. From Multi-Channel to "Omnichannel" Retailing: Review of the Literature and Calls for Research. In Proceedings of the 2nd International Conference on Contemporary Marketing Issues, Athens, Greece, 18–20 June 2014; pp. 18–20.

50. Koh, S.L.; Gunasekaran, A.A. knowledge management approach for managing uncertainty in manufacturing. *Ind. Manag. Data Syst.* **2006**, *106*, 439–459. [CrossRef]

51. Chow, H.K.; Choy, K.L.; Lee, W.B.; Chan, F.T. Integration of web-based and RFID technology in visualizing logistics operations—A case study. *Supply Chain Manag.* **2017**, *12*, 221–234. [CrossRef]

52. Lai, F.; Li, D.; Wang, Q.; Zhao, X. The information technology capability of third-party logistics providers: A resource-based view and empirical evidence from China. *J. Supply Chain Manag.* **2008**, *44*, 22–38. [CrossRef]

53. Pokharel, S. Perception on information and communication technology perspectives in logistics: A study of transportation and warehouses sectors in Singapore. *J. Enterp. Inf. Manag.* **2005**, *18*, 136–149. [CrossRef]

54. Giaglis, G.M.; Minis, I.; Tatarakis, A.; Zeimpekis, V. Minimizing logistics risk through real-time vehicle routing and mobile technologies: Research to date and future trends. *Int. J. Phys. Distrib. Logist.* **2004**, *34*, 749–764. [CrossRef]

55. Oláh, J.; Karmazin, G.; Pető, K.; Popp, J. Information technology developments of logistics service providers in Hungary. *Int. J. Logist. Res. Appl.* **2018**, *21*, 332–344. [CrossRef]

56. Wang, Q.; Lai, F.; Zhao, X. The impact of information technology on the financial performance of third-party logistics firms in China. *Supply Chain Manag.* **2008**, *13*, 138–150. [CrossRef]

57. Wu, F.; Yeniyurt, S.; Kim, D.; Cavusgil, S.T. The impact of information technology on supply chain capabilities and firm performance: A resource-based view. *Ind. Mark. Manag.* **2006**, *35*, 493–504. [CrossRef]

58. Colla, E.; Dupuis, M. Research and managerial issues on global retail competition: Carrefour/Wal-Mart. *Int. J. Retail Distrib. Manag.* **2002**, *30*, 103–111. [CrossRef]

59. Vowels, S.A. A strategic case for RFID: An examination of Wal-Mart and its supply chain. In Proceedings of the Ninth Annual Conference of the Southern Association for Information Systems, Jacksonville, FL, USA, 8–9 December 2006; pp. 148–152.

60. Kim, C.; Yang, K.H.; Kim, J. A strategy for third-party logistics systems: A case analysis using the blue ocean strategy. *Omega* **2008**, *36*, 522–534. [CrossRef]

61. Lieb, R.C.; Lieb, K.J. Is Amazon a 3PL. *Supply Chain Quarterly*; Quarter 3 2014. Available online: http://www.supplychainquarterly.com/topics/Logistics/20141027-is-amazon-a-3pl/ (accessed on 26 July 2018).

62. Busse, C. A procedure for secondary data analysis: Innovation by logistics service providers. *J. Supply Chain Manag.* **2010**, *46*, 44–58. [CrossRef]

63. Ling, T.K.; Lee, C.K.; Ho, W. The analysis and case studies of successful express logistics companies. *Int. J. Value Chain Manag.* **2009**, *3*, 20–35. [CrossRef]

64. Stock, J.R. Doctoral Research in logistics and logistics-related areas: 1992–1998. *J. Bus. Logist.* **2001**, *22*, 125–256. [CrossRef]

65. Tseng, Y.H.; Lin, C.J.; Lin, Y.I. Text mining techniques for patent analysis. *Inf. Process. Manag.* **2007**, *43*, 1216–1247. [CrossRef]

66. Archibugi, D.; Planta, M. Measuring technological change through patents and innovation surveys. *Technovation* **1996**, *16*, 451–519. [CrossRef]

67. Ernst, H. The use of patent data for technological forecasting: The diffusion of CNC-technology in the machine tool industry. *Small Bus. Econ.* **1997**, *9*, 361–381. [CrossRef]

68. Narin, F. Patent bibliometrics. *Scientometrics* **1994**, *30*, 147–155. [CrossRef]

69. Narin, F.; Noma, E.; Perry, R. Patents as indicators of corporate technological strength. *Res. Policy* **1987**, *16*, 143–155. [CrossRef]

70. Lei, X.P.; Zhao, Z.Y.; Zhang, X.; Chen, D.Z.; Huang, M.H.; Zheng, J.; Liu, R.S.; Zhang, J.; Zhao, Y.H. Technological collaboration patterns in solar cell industry based on patent inventors and assignees analysis. *Scientometrics* **2013**, *96*, 427–441. [CrossRef]

71.　Choi, J.Y.; Jeong, S.; Kim, K. A study on diffusion pattern of technology convergence: Patent analysis for Korea. *Sustainability* **2015**, *7*, 11546–11569. [CrossRef]

72.　Yun, J.J.; Jeong, E.; Park, J. Network analysis of open innovation. *Sustainability* **2016**, *8*, 729. [CrossRef]

73.　Cho, T.S.; Shih, H.Y. Patent citation network analysis of core and emerging technologies in Taiwan: 1997–2008. *Scientometrics* **2011**, *89*, 795. [CrossRef]

74.　Kim, G.; Lee, J.; Jang, D.; Park, S. Technology clusters exploration for patent portfolio through patent abstract analysis. *Sustainability* **2016**, *8*, 1252. [CrossRef]

75.　Yoon, B.; Park, Y. A text-mining-based patent network: Analytical tool for high-technology trend. *J. High Technol. Manag. Res.* **2004**, *15*, 37–50. [CrossRef]

76.　Chen, H.; Zhang, G.; Zhu, D.; Lu, J. Topic-based technological forecasting based on patent data: A case study of Australian patents from 2000 to 2014. *Technol. Forecast. Soc.* **2017**, *119*, 39–52. [CrossRef]

77.　Jeong, B.; Yoon, J. Competitive intelligence analysis of augmented reality technology using patent information. *Sustainability* **2017**, *9*, 497. [CrossRef]

78.　Lee, W.S.; Han, E.J.; Sohn, S.Y. Predicting the pattern of technology convergence using big-data technology on large-scale triadic patents. *Technol. Forecast. Soc.* **2005**, *100*, 317–329. [CrossRef]

79.　Kuusi, O.; Meyer, M. Anticipating technological breakthroughs: Using bibliographic coupling to explore the nanotubes paradigm. *Scientometrics* **2007**, *70*, 759–777. [CrossRef]

80.　Van Merkerk, R.O.; Van Lente, H. Tracing emerging irreversibilities in emerging technologies: The case of nanotubes. *Technol. Forecast. Soc.* **2005**, *72*, 1094–1111. [CrossRef]

81.　Suominen, A.; Toivanen, H.; Seppänen, M. Firms' knowledge profiles: Mapping patent data with unsupervised learning. *Technol. Forecast. Soc.* **2017**, *115*, 131–142. [CrossRef]

82.　Alghamdi, R.; Alfalqi, K. A survey of topic modeling in text mining. *Int. J. Adv. Comput. Sci. Appl.* **2015**, *6*, 147–153. [CrossRef]

83.　Blei, D.M. Probabilistic topic models. *Commun. ACM* **2012**, *55*, 77–84. [CrossRef]

84.　Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.

85.　Steyvers, M.; Griffiths, T. Probabilistic topic models. In *Handbook of Latent Semantic Analysis*; Lawrence Erlbaum Associates, Inc.: Mahwah, NJ, USA, 2007; pp. 424–440.

86.　Nigam, K.; McCallum, A.K.; Thrun, S.; Mitchell, T. Text classification from labeled and unlabeled documents using EM. *Mach. Learn.* **2000**, *39*, 103–134. [CrossRef]

87.　Grün, B.; Hornik, K. Topicmodels: An R package for fitting topic models. *J. Stat. Softw.* **2011**, *40*, 1–30. [CrossRef]

88.　Hofmann, T. Probabilistic Latent Semantic Indexing. In *ACM SIGIR Forum*; ACM: New York, NY, USA, 2017; Volume 51, pp. 211–218.

89.　Kim, J.; Lee, S. Patent databases for innovation studies: A comparative analysis of USPTO, EPO, JPO and KIPO. *Technol. Forecast. Soc.* **2015**, *92*, 332–345. [CrossRef]

90.　Schmoch, U. Concept of A Technology Classification for Country Comparisons. Final Report to the World Intellectual Property Organisation (WIPO). 2008. Available online: http://www.wipo.int/export/sites/www/ipstats/en/statistics/patents/pdf/wipo_ipc_technology.pdf (accessed on 3 June 2008).

91.　Lummus, R.R.; Krumwiede, D.W.; Vokurka, R.J. The relationship of logistics to supply chain management: Developing a common industry definition. *Ind. Manag. Data Syst.* **2001**, *101*, 426–432. [CrossRef]

92.　Cooper, M.C.; Lambert, D.M.; Pagh, J.D. Supply chain management: More than a new name for logistics. *Int. J. Logist. Manag.* **1997**, *8*, 1–14. [CrossRef]

93.　Giunipero, L.C.; Brand, R.R. Purchasing's role in supply chain management. *Int. J. Logist. Manag.* **1996**, *7*, 29–38. [CrossRef]

94.　Feldman, R.; Sanger, J. *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*; Cambridge University Press: Cambridge, UK, 2007.

95.　Hotho, A.; Nürnberger, A.; Paaß, G. A Brief Survey of Text Mining. In *Ldv Forum*; AIS: Bangkok, Thailand, 2005; Volume 20, pp. 19–62.

96.　Kuang, Q.; Xu, X. Improvement and Application of TF IDF Method Based on Text Classification. In Proceedings of the 2010 International Conference on Internet Technology and Applications, Wuhan, China, 8 August 2010; pp. 1–4.

97.　Zhang, W.; Yoshida, T.; Tang, X. A comparative study of TF* IDF, LSI and multi-words for text classification. *Expert. Syst. Appl.* **2011**, *38*, 2758–2765. [CrossRef]

98. Ernst, H. Patent information for strategic technology management. *World Patent Inf.* **2003**, *25*, 233–242. [CrossRef]

99. Fabry, B.; Ernst, H.; Langholz, J.; Köster, M. Patent portfolio analysis as a useful tool for identifying R&D and business opportunities—An empirical application in the nutrition and health industry. *World Patent Inf.* **2006**, *28*, 215–225.

100. Deng, Z.; Lev, B.; Narin, F. Science and technology as predictors of stock performance. *Financ. Anal. J.* **1999**, *55*, 20–32. [CrossRef]

101. Ernst, H. Patenting strategies in the German mechanical engineering industry and their relationship to company performance. *Technovation* **1995**, *15*, 225–240. [CrossRef]

102. Hall, B.H.; Jaffe, A.B.; Trajtenberg, M. Market Value and Patent Citations: A First Look. NBER Working Paper 7741. National Bureau of Economic Research. 2000. Available online: http://www.nber.org/papers/w7741.pdf (accessed on 7 March 2018).

103. Lerner, J. The importance of patent scope: An empirical analysis. *Rand. J. Econ.* **1994**, 319–333. [CrossRef]

104. Shane, S. Technological opportunities and new firm creation. *Manag. Sci.* **2001**, *47*, 205–220. [CrossRef]

105. Jaffe, A.B.; Trajtenberg, M. *Patents, Citations, and Innovations: A Window on the Knowledge Economy*; MIT Press: Cambridge, MA, USA, 2002.

106. Sauvage, T. The relationship between technology and logistics third-party providers. *Int. J. Phys. Distrib. Logist.* **2003**, *33*, 236–253. [CrossRef]

107. United States Patent and Trademark Office. Overview of the U.S. Patent Classification System (USPC). 2012. Available online: https://www.uspto.gov/sites/default/files/patents/resources/classification/overview.pdf (accessed on 7 March 2018).

108. Phan, X.H.; Nguyen, L.M.; Horiguchi, S. Learning to Classify Short and Sparse Text & Web with Hidden Topics from Large-Scale Data Collections. In Proceedings of the 17th International Conference on World Wide Web, Beijing, China, 21–25 April 2008; ACM: New York, NY, USA; pp. 91–100.

109. Ovans, A. Can you patent your business model. *Harv. Bus. Rev.* **2000**, *78*, 16.

110. United States Patent and Trademark Office. Automated Financial or Management Data Processing Methods (Business Methods). USPTO White Paper. 2000. Available online: https://www.uspto.gov/sites/default/files/web/menu/busmethp/whitepaper.pdf (accessed on 7 March 2018).

111. Wengler, S.; Biermann, P. Logistics Startups Are Booming. *SIMON·KUCHER&PARTNERS*, 2 March 2018. Available online: https://www.simon-kucher.com/en-be/blog/logistics-startups-are-booming (accessed on 6 June 2018).

112. CB Insights. Amazon's 'Beehive', Drone-Carrying Trains Reinforce Focus on Logistics Tech. 2017. Available online: https://www.cbinsights.com/research/amazon-warehouse-patent/ (accessed on 3 March 2018).

113. Garrison, G. An assessment of organizational size and sense and response capability on the early adoption of disruptive technology. *Comput. Hum. Behav.* **2009**, *25*, 444–449. [CrossRef]

114. Green, S.G.; Gavin, M.B.; Aiman-Smith, L. Assessing a multidimensional measure of radical technological innovation. *IEEE Trans. Eng. Manag.* **1995**, *42*, 203–214. [CrossRef]

115. Foster, R.N. Timing technological transitions. *Technol. Soc.* **1985**, *7*, 127–141. [CrossRef]

116. Lee, P.C.; Su, H.N. How to analyze technology life cycle from the perspective of patent characteristics? In Proceedings of the 2015 Portland International Conference on Management of Engineering and Technology (PICMET), Portland, OR, USA, 8 August 2015; pp. 2079–2083.

117. Dahlin, K.B.; Behrens, D.M. When is an invention really radical? Defining and measuring technological radicalness. *Res. Policy* **2005**, *34*, 717–737. [CrossRef]