Linking Information Technology and Entrepreneurship: A Literature Review

Yousra Harb  
Yarmouk University, Irbid, Jordan  
yousra.harb@yu.edu.jo

Yanyan Shang  
The University of Tampa, Florida, USA  
yshang@ut.edu

Abstract
This study reviews research carried out in the domain of IT and entrepreneurship. A total of 907 papers, published between 1978 and 2020 were used to uncover the latent topics addressed in the domain. A topic modeling (LDA) algorithm was used to automate the process of extracting the initial research topics from the data. The literature review further enhances the understanding of IT-associated entrepreneurship research, providing useful insights for future research and informs practice in this domain of study.  

Keywords: IT, Entrepreneurship, Topic modeling, Literature review

1. Introduction

Information technology is considered as one of the most essential drivers of economic and social value and plays an indispensable role in transforming organizations, markets, and industries [1]. Waves of technological changes create new opportunities that need to be leveraged and exploited [2]. Entrepreneurs are one of the change agents in terms of exploiting the new opportunities brought by the new technologies such as big data, business intelligence, cloud computing, augmented reality, mobile services, or internet of things [2], [3].

The reasons for adopting IT are to achieve growth, stay competitive and enhance innovation abilities [4]. The business model is one of the main tools for absorbing technological perspectives of innovations. Particularly, entrepreneurs design and implement new business models through the enabling of information technology capabilities [5]. New technologies also provide means to enhance information and knowledge capture and sharing, lower the cost of production and labor, and add value to products and services and thus enhance the quality [4]. Administration processes can be also technology intensive areas in businesses to enhance performance [1]. By adopting IT and digital technologies also initiates new ways of doing business such as digital business models and digital entrepreneurship [3].

Moreover, in this context, IT was also proven to influence the organizational design in terms of job design through supporting the autonomy within the organization and thus influencing employees’ attitudes toward involvement in entrepreneurial activities; lateral linkage design concerns with enhancing analytical and design capabilities; and decision-making design by taking the advantage of both centralization and decentralization design [1].

Due to the importance of information technology to entrepreneurs and the great advantages it offers them; it has received tremendous attention from various discipline areas. Research linking information technology and entrepreneurship fields is needed in order to improve the understanding of the role IT plays in this young field. On a related note, a review of literature relating to IT and entrepreneurship can provide insights and inform research in the domain. However, limited research has reviewed the studies that have investigated these two fields together. Steininger [2] conducted a systematic review of information systems, entrepreneurship, and general and strategic management literature using 292 articles to provide an overview of IT and entrepreneurship research.

Other research studies have more specific focus. For instance, Zaheer et al. [6] study provided a structured review of digital entrepreneurship literature by analyzing 133 articles in this discipline. Olanrewaju and Hossain [7] systematically reviewed 160 articles published between 2002 and 2018 in social media and entrepreneurship domain. However, our study is the first attempt to survey a large corpus of IT and entrepreneurship-related articles. In addition, this study provides an understanding of the main research topics of IT-associated entrepreneurship research using topic modeling algorithm.
The purpose of this study is to uncover the directions of IT-associated entrepreneurship by analyzing the abstracts listed in 907 publications from 1978 to 2020. Several research studies reviewed literature in various domains using articles metadata. In management information systems, Delen and Crossland [8] applied text mining approach to analyze the titles and abstracts of 1123 articles. Abu-Shanab and Harb [9] conducted text mining on articles keywords in e-government area of research. Bragge et al. [10], analyzed articles’ keywords using text mining in multiple criteria decision-making area of research. In the field of operation research and management science, Gatti et al. [11] used LDA to uncover the latent topics of abstracts from 37 related journals. Worth noting, abstract data contains a high density of words, therefore making it more appropriate for LDA algorithm [11].

The organization of the paper is as follows: section 2 discusses the research method, section 3 presents the results. Analysis is discussed in section 4, and finally the conclusion is outlined in section 5.

2. Method

According to Webster and Watson [12], a literature review of prior studies is a critical step to analyze a given subject and synthesize previous research. Such approach would strengthen a particular field of study and provide important clues on current and future research directions. In fact, it is an essential approach to identify trends of research that can unveil new studies and discoveries [13].

The main research question of this research involves identifying the key IT/IS and entrepreneurship research topics and research methods witnessed in the dataset (from 1978 to 2020). Basically, our objective is to use an automated text mining to analyze prior studies in IT/IS and entrepreneurship and understand the current status of this interdisciplinary topic and provide a snapshot to inform future research. Given the high volume of research articles returned by hitting online databases, reading and open coding the content of each paper is a challenging and intensive labor task [14]. Therefore, automating the process of reading and open coding the content enable the research to extract useful topics from large data without being contaminated by any preconceptions [15].

The process of conducting this review is described in the next section.

2.1 Data collection

Interrelating information technology field with entrepreneurship context is interdisciplinary topic and requires cross-domain in the areas of information technology, information systems and entrepreneurship to investigate the use of information technology for business management, three prominent databases include business management journals were used to source literature; Web of Science, Scopus, and ABI. We included only refereed journal articles in the search result. This approach is consistent with previous studies on entrepreneurship that have include only journal articles (e.g., [7]).

To investigate IT and entrepreneurship, first we used the keywords (“information techno*”, and “entrepreneur*”). However, we noticed that studies in this domain also use other IT-related terms such as information systems, information communication technology (ICT), or computerized information. Therefore, we included keywords reflecting information technology term. The final search query was: 

("entrepreneur*" AND "computerized information") 
OR ("entrepreneur*" AND "Information system*) 
OR ("entrepreneur*" AND "information technolog") 
OR ("entrepreneur*" AND "information and communication technology") 
OR ("entrepreneur*" AND "computer technology") 
OR ("entrepreneur*" AND infotech) 
OR ("entrepreneur*" AND "computer-based informat") 
OR ("entrepreneur*" AND "computer-human")

We limited the search to the abstract to have focused results and retrieve relevant articles. As aforementioned, only journal articles published in English were selected. In line with prior systematic review studies, we excluded articles published in conferences, books, book chapters, theses, and reports [7]. This resulted in 408 hits in Web of science, 596 hits in Scopus, and 305 articles from ABI database. Through manual screening and after removing duplicate articles the shortlist contained 907 articles. The time period under investigation is between 1978 and 2020. Figure 1 presents the distribution of articles by year.
2.2 Topic modeling: LDA

As the number of articles is large, extracting useful information by manual coding might be a time-consuming and a tedious task [16]. Hence, text mining was utilized to facilitate the process of extracting useful topics from the literature. In this approach, we adopted topic modeling (LDA) algorithm to unveil hidden topics from a collection of documents [17]. LDA is a generative probabilistic topic modeling algorithm that outperforms over other well-known topic models.

The preprocessing of the data included the following transformations: (1) we removed the stop words and we also extends the stop words to include some other noisy words such as ('among', 'use', 'subject', 'edu', 'use', 'every', 'also', 'one'), (2) lemmatization was performed on the documents, and (3) The documents were also converted to lower case. We selected the Latent Dirichlet Allocation (LDA) model from Gensim package available in Python software as one of the most important probabilistic used algorithms for topic identification [14]. LDA is a generative probabilistic topic models from a collection of documents. In this generative process, each document is modeled as a multinomial distribution over k topics and it models each topic as a multinomial distribution over vocabulary V.

Choosing the right number of topics in LDA is an important task. Literature has proposed some measures to evaluate the quality of the generated topics such as predictive likelihood and coherence measures [18], [19]. Past studies have shown the predictive likelihood may not yield interpretable topics from human judgment [19]. Moreover, when the generated topics are used to understand the trends or the development of a particular field of study, it is important to apply measures that serve this purpose [19]. Thus, to serve the motivation of this study pertains to understanding the development of IT/IS and entrepreneurship field, we used topic coherence measures to uncover the coherence of a topic. Each generated topic consisted of words, and the topic coherence measures score a topic by measuring the semantic similarity of the top N words in the topic.

2.3 Creating LDA models

The following procedures were followed to create LDA models in this study: (1) we built the base LDA model with 10 topics where each topic is a combination of words and a certain weight associated with each word, (2) the baseline coherence score was calculated for the baseline model and the value was (0.32), (3) sensitivity tests were then performed to determine the model hyperparameters [number of topics (k), Dirichlet hyperparameter alpha, and Dirichlet hyperparameter beta]. These tests were performed in sequence over a validation corpus set, (4) for every LDA model created (541 in total), the coherence score was calculated. To determine the optimal number of topics, we investigated the coherence scores corresponds to varying number of topics. In line with the previous studies (e.g., [18]), we selected the model that yielded the highest coherence score (See Figure 2). In this case, we chose k=16. The values of alpha and beta parameters were selected based on the maximum coherence score for k=16. We then trained the final model using the selected parameters. The coherence score for the final model was 0.46 (14% improvement over the baseline model).

![Figure 1. Distribution of articles by year](image1)

![Figure 2. Calculated CV topic coherence score for LDA models with K = [2, ..., 20]](image2)
3. Results

This section reports and discusses the results of the LDA topic modeling. The LDA topic modeling conducted in the previous section resulted in a list of top words, whose latent meanings can be regarded as topics. A list of top words per topic is provided, as well as a score that indicates the weighted presence of the top word per topic (topic loading). Table 1 shows the top words generated from the topic modeling and the topics that are extracted from the titles and abstracts of the selected literature. The consensus about the meaning of the 16 topics was reached during meetings with the co-authors.

<table>
<thead>
<tr>
<th>No.</th>
<th>Key words</th>
<th>Topic</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>institutional, transformation, sustainability, toward, pathway, lock, unsustainable, tipping, innovation, scale</td>
<td>Institutional innovation and transformation</td>
<td>1.40%</td>
</tr>
<tr>
<td>2</td>
<td>safety, management, practice, policy, health, commitment, influence, significantly, training, worker</td>
<td>Organizational environment and commitment</td>
<td>2.10%</td>
</tr>
<tr>
<td>3</td>
<td>business, information, entrepreneurship, entrepreneur, development, failure, process, market, technology, firm</td>
<td>Business process management</td>
<td>18%</td>
</tr>
<tr>
<td>4</td>
<td>job, self, esteem, decision, founding, intention, attitude, career, capability, cognitive</td>
<td>Qualification/ characteristics of entrepreneurs</td>
<td>2.30%</td>
</tr>
<tr>
<td>5</td>
<td>craft, marketing, project, sector, hotel, quality, policy, leadership, success, artisan</td>
<td>Policies and industries</td>
<td>2.80%</td>
</tr>
<tr>
<td>6</td>
<td>entrepreneur, migrant, work, among, infotech, group, experience, access, issue, high</td>
<td>Labor market and migrant</td>
<td>2.30%</td>
</tr>
<tr>
<td>7</td>
<td>job, management, entrepreneurial, industry, sector, resource, information, indian, mobility, npd</td>
<td>Industries and resource</td>
<td>7.30%</td>
</tr>
<tr>
<td>8</td>
<td>product, catfish, pantura, problem, woman, several, material, raw, good, cost</td>
<td>Operational management</td>
<td>2.30%</td>
</tr>
<tr>
<td>9</td>
<td>information, technology, business, model, entrepreneur, system, development, service, based, new</td>
<td>Business model and service</td>
<td>34.50%</td>
</tr>
<tr>
<td>10</td>
<td>technology, consumer, information, entrepreneurship, management, right, new, development, entrepreneurial, product</td>
<td>Customer service</td>
<td>0.80%</td>
</tr>
<tr>
<td>11</td>
<td>woman, university, development, education, entrepreneurship, job, information, remote, modern, ict</td>
<td>Woman/Education</td>
<td>5.10%</td>
</tr>
<tr>
<td>12</td>
<td>network, performance, rural, ipo, smart, technology, firm, government, analysis, entrepreneurial</td>
<td>Initiative and Technology adoption</td>
<td>6%</td>
</tr>
<tr>
<td>13</td>
<td>social, firm, start, digital, new, enterprise, effect, capital, survival, entrepreneurship</td>
<td>Social impact</td>
<td>5.80%</td>
</tr>
<tr>
<td>14</td>
<td>entrepreneurship, cluster, treatment, relationship, health, component, implementation, sud, software</td>
<td>Entrepreneurial strategic orientation</td>
<td>3.30%</td>
</tr>
<tr>
<td>15</td>
<td>start, internet, ups, founder, early, social, stage, capital, success, model</td>
<td>Initiative</td>
<td>1.40%</td>
</tr>
<tr>
<td>16</td>
<td>accounting, discourse, management, student, organisation, waste, within, organisational, school, information</td>
<td>Administrative entrepreneurship</td>
<td>4.50%</td>
</tr>
</tbody>
</table>

16 topics discovered from the topic modeling were first mapped to 8 aspects of IT-associated entrepreneurship research. The mappings between the topics and the aspects of IT-associated entrepreneurship research can be one-to-one. For example, the topics “Business process management”,...
“Innovation and transformation”, “Leadership”, and “Operations management” were mapped into their own categories. There are also multiple topics that are mapped to the same category. For example, “Policies and industries”, “Industries and resource”, and “Woman/Education” were mapped into the same aspect of entrepreneurship “Industry analysis”. Table 2 summarizes how the topics mapped into 8 aspects of IT-associated entrepreneurship.

Table 2. The aspects of IT-associated entrepreneurship discovered from topic modeling

<table>
<thead>
<tr>
<th>Aspect of IT-associated Entrepreneurship</th>
<th>Topics</th>
<th>Weight Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business process management</td>
<td>Business process management</td>
<td>18%</td>
</tr>
<tr>
<td>Industry analysis</td>
<td>Industries and resource</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>Policies and industries</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Woman/Education</td>
<td></td>
</tr>
<tr>
<td>Initiative</td>
<td>Initiative</td>
<td>7%</td>
</tr>
<tr>
<td></td>
<td>Initiative and Technology adoption</td>
<td></td>
</tr>
<tr>
<td>Transformation</td>
<td>Institutional innovation and transformation</td>
<td>1%</td>
</tr>
<tr>
<td>Leadership</td>
<td>Qualification/characteristics of entrepreneurs</td>
<td>2%</td>
</tr>
<tr>
<td>Operations management</td>
<td>Operations management</td>
<td>2%</td>
</tr>
<tr>
<td>Organizational management</td>
<td>Administrative entrepreneurship</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>Labor market and migrant</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Organizational environment and commitment</td>
<td></td>
</tr>
<tr>
<td>Strategy</td>
<td>Business model and service</td>
<td>44%</td>
</tr>
<tr>
<td></td>
<td>Customer service</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Entrepreneurial strategic orientation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Social impact</td>
<td></td>
</tr>
</tbody>
</table>

4. Analysis

This section discusses the findings and implications of our study. As shown in Figure 3, Strategy, business process management, and industry analysis are the most well-studied areas in IT-associated entrepreneurship literature, whereas transformation, leadership, and Operations management are not well-covered areas of research in the domain.

Figure 3. The 8 aspects of IT-associated entrepreneurship discovered from topic modeling

Business process management and Operations management

Business process management (BPM) combines knowledge from information technology and management sciences and applies it to operational business processes [20], [21]. BPM and Operations Management (OM) are closely related to organizations’
productivity and cost control, and thereby, has received considerable attentions. Studies with regarding to BPM and OM focus on the role of information technology adoption in business process improvement. For example, Sacco and Strait [22] explored the adoption of IT infrastructure to facilitate the transition of a nonprofit organization to pursue an entrepreneurial path. Wu et al. [23] conducted an empirical study to investigate the factors that affect the organizations’ propensity to adopt cloud computing technologies.

We suggest that future studies interested in this area can investigate how to utilize information technology to facilitate and improve the process of entrepreneurship itself. Entrepreneurship is a journey which starts from a new creative idea and a whole process to convert the idea into a business. information technology can be adopted not only to improve the business model, but also the entire journey of converting the business idea into an enterprise.

*Industry analysis*

Industry-level research was a popular topic in this area. Articles related to this topic investigated the role of IT as a central part for value creation. By exploring the dataset in this study, we found that researchers studied IT-associated entrepreneurship related themes in several industries such as education, healthcare, tourism and hospitality, IT and software industry, manufacturing industries, agriculture industry, banking industry, telecommunication, construction industry, and transportation industry. Other investigated domains include, government, financial, service, food, airline, music, fashion, and semiconductor. We also noticed that IT-associated entrepreneurship related research areas in education were well-covered sector. Tourism and hospitality, Healthcare, and software and IT industries are also well-studied industries in this domain of research.

*Initiative*

Entrepreneurship initiative refer to the ability to turn ideas into action. “It includes the ability to plan and manage projects in order to achieve objectives, but also creativity, innovation, and risk taking” [24]. There are general two ways to initiate an entrepreneurial project: creating new ideas or adopting technologies or methods for imitating the new combination of production already available [25]. We found that literature in this area was largely interested both in the adoption of information technology (IT) and information and communication technology (ICT). For example, studies [26]–[29] investigated the impact of IT/ICT on entrepreneurship and economies in developing countries (Africa and Asia) or rural regions in Europe.

However, less studies have focused on the new role of information technology in the process of initiating or innovating. We suggest that future studies in this area can develop frameworks to show how information technology can be helpful of incubating entrepreneurial initiation such as capturing new ideas and visualizing or prototyping new ideas.

*Transformation*

Studies under this area of research concerns with investigating the transformational or disruptive role of IT in business. Companies may need to transform themselves in significant ways to succeed in this digital world. Moreover, research studies addressed the role of IT plays in transforming both innovation and entrepreneurship. Particularly, IT catalysts new forms of innovation, create new business models, new customer experiences, and new types of products and services [30].

As shown in Figure 3, we noticed this area is under explored. Hence, future research may further investigate this interdisciplinary area to provide a more comprehensive analysis about the implications of transformation of innovation and entrepreneurship enabled by new technologies on individual, organizations, and industries.

*Leadership*

Topics covered under this research area include adoption of ICT by entrepreneurs. The tendency to adopt IT is highly influenced by the owner of the company. Research emphasizes the characteristics of the leadership such as innovation and behavior toward IT adoption [31]. Top managers entrepreneurial leadership is recognized as essential in perceiving IT innovativeness in organizations. They are the critical facilitating factor to IT innovativeness [32]. Future studies in this area can further investigate the impact of entrepreneurial leadership style and characteristics on creativity and innovation.

*Organizational management*

Organizational management plays a crucial role once the ideas have been turning into actions. Studies in this area focused on organizational behavior including the administrative actions of handling routine tasks, labor market, and organizational environment such as employee safety, satisfaction, and commitment. Some studies investigated the impact of different demographic character of the entrepreneurs, such as the immigration status and the experiences [33], [34], the gender differences [35], while other studies looked into organization’s success by examining work-related injuries and regulations [36], and job performance and organizational commitment [37].
While organizational management is rather behavior than technical, IT may affect employee’s problem solving and decision-making process. We suggest that studies can be conducted from the perspective how IT improves work environment and how IT enables fast information and knowledge sharing for better solutions and decisions.

**Strategy**

Strategy refers to a plan of action or policy designed to achieve a major or overall aim. About 44% of selected literature investigates areas about strategical IT-associated entrepreneurship, for instance, the strategic orientation, the business models, the customer service, and the social capital and impact. Entrepreneurship scholars have genuinely attempted to understand what opportunities to bring into existence, and how future goods and services are discovered and exploited to create and grow new ventures [38]. Studies in this area investigated activities in different business functional units in the value chain to understand how these activities are pertinent to their organizational strategies effectively. The strategic entrepreneurship has been highly of interests to scholars and many excellent studies have been produced. We found that most studies in this area focused on economic opportunities. We suggest that future study may consider investigating non-profit organizations and/or the philanthropic responsibilities of an entrepreneurial company.

5. Conclusion

A literature review is useful and helpful in examining how well an area of research has been studied. However, traditional ways of conducting literature review involving a time-consuming manual process and a large amount of subjective reading. In this study, we proposed to use a semi-auto method to perform a literature review. The LDA topic modeling can mine the topics quickly and fairly in a relatively objective and scientific way. We believe it can help researchers to perform a thorough yet less time-consuming literature review. Further, we utilized the LDA topic modeling to analyze 907 literature in an interdisciplinary area IT-associated entrepreneurship. Our findings suggested several insights on future studies in the area.

Our study is not without limitation. First, this study is limited by the sample selection. The focus was primarily on the journal articles. Conferences papers, books, and other journal articles published in other databases were not included in the analysis. Therefore, the extracted topics in IT-associated entrepreneurship interdisciplinary field might not reflect the topics in the related publications that were excluded from this study. Second, this study analyzed part of the articles metadata. However, in some cases abstract do not reflect all the content of the paper. Hence, future research can extend this study by employing full text analysis.

6. References


