Claudiu CICEA, PhD E-mail: claudiu.cicea@man.ase.ro The Bucharest University of Economic Studies Tsoulfidis LEFTERIS, PhD Email: Int@uom.edu.gr University of Macedonia Corina MARINESCU, PhD E-mail: corina.marinescu@man.ase.ro The Bucharest University of Economic Studies Ștefan Cătălin POPA, PhD Student E-mail: popacatalinstefan@yahoo.com The Bucharest University of Economic Studies Cătălina Florentina ALBU, PhD Student E-mail: albucatalinaflorentina@yahoo.com

APPLYING TEXT MINING TECHNIQUE ON INNOVATION-DEVELOPMENT RELATIONSHIP: A JOINT RESEARCH AGENDA

Abstract. Throughout the time, innovation gained its place among factors influencing development. To identify the potential relationships between innovation and development and also the myriad of syntaxes related to their intersection in the literature, the authors create a joint research agenda by analyzing conclusive contents, based on the text mining technique. The research methodology follows three steps, including: selecting articles, creating the dictionary (with main categories and subcategories) and text mining which provides insights on frequencies of occurrence for dictionary categories, clusters formed among main discussed subjects, the existence of centroids within clusters and different types of ties among syntaxes. The main findings are grouped on three time intervals. The study's results are discussed according to four dimensions (time, economic, geographic, semantic) seen as different aspects that can help shaping an overview on research in innovation and development.

Keywords: Development, Innovation, Text mining.

JEL Classification: O30, O50

1. Introduction

In the context of a globalized economy, researchers from different countries have tried to find out as much as possible about the factors affecting economic growth. Over time, key concepts have been analyzed to explain their influence on the variability of economic growth. One of these is innovation, which is reflected by the Schumpeterian models. As an example, we have the quality scale model of Aghion and Howitt (1992), defined as "a model of endogenous growth ... in which vertical innovations, generated by a competitive research sector, constitute the underlying source of growth".

Over time, the concept of innovation has gained the attention and interest of many researchers. The diversity of issues addressed in relation to innovation generates an assortment of papers and articles published over the past decades. Those issues refer to: the variety of areas in which the researchers deal with the innovation concept (Bagchi-Sen & Scully, 2004; Kozak, 2013; Marinescu, Ciocoiu & Cicea, 2015): bioenergy, biotechnology, tourism, waste and recycling, high-tech industry, management or to the geographic prevalence of analyses (Aylen, 2010; Potts, 2010; Horbach, 2016): Europe, Asia, North America, Africa, Australia. The innovation type also divided articles after (Reveiu & Dardala, 2013; Cicea, Marinescu & Moroianu, 2015; Feng & Chen, 2018): product innovation, green innovation, process innovation, disruptive innovation.

Therefore, different relationships can be identified in the innovation – development syntax. The first one is innovation development, which refers to the development of innovation as a continuous, universal activity that occurs in the economy (Hudson & Minea, 2013), but also to the development of a certain type of innovation, for instance: development of industrial innovation, development of environmental innovation, development of strategy innovation (Stankevice & Jucevicius, 2013) and so on. The second one is innovation for development, which refers to the multitude of innovation elements and activities that contribute to the development of a field (Feng & Chen, 2018). The third one is innovation and R&D, seen as two elements that may influence the performance in one particular area (Hall, 2002) or which may impact each other (Fritsch & Franke, 2004). The last one is R&D innovation, which refers to activities of innovation in research and development field (Allahyari et al., 2017).

In the following we aim to analyze the relationship between innovation and development, as it was studied by researchers in various specialized journals within the numerous articles treating the concerned relationship in different forms. The originality of our work consists in combining both qualitative and quantitative analysis of a large sample of articles in order to highlight foundations, trends and specific characteristics of innovation-development research field. The proposed methodology represents an innovative idea for studying bibliographic material and exploring scholarly contents. The research is organized as follows: at first we

explain the research methodology, then we identify and analyze in detail the main results and in the end we draw conclusions.

2. Materials and Methods

The approach chosen for the analysis is Text Mining, to identify the potential relationships between innovation and development. In a previous research, the authors used bibliometric analysis in order to identify scientific interest points and trends in a specific period of time (Cicea & Marinescu, 2019). The purpose of applying this technique is to obtain valuable information about the concerned subject. Also called knowledge discovery from text, text mining represents a complex process which needs a data structure for the analyzed text (Smeureanu & Bucur, 2012; Cernian et al., 2016; Shi et al., 2018). For instance, our analyzed texts have titles, authors' names, abstracts, keywords, introductions and so on. The methodology involves going through several stages, as follows: 2.1. The selection of articles to be analysed

Regarding the selection of articles, it was based on Bradford's law and the law of Garfield (1980). Bradford's law, which appeared in 1934, is described by Garfield himself in one of his works. Given the large number of articles analyzing innovation in relation to development and the variety of areas in which it occurs, the final result of the selection is to create a representative sample of articles, journals and publishers for the entire analyzed field. Table 1 includes, according to each editor, the journals, along with the number of articles provided for this analysis.

The following criteria were useful to identify the articles, the journals, and the publishers that became the subject of the analysis. Firstly, the Scopus and Web of Science databases were used. These two databases are among the most well-known databases for indexing articles and among the most widely used by researchers, providing access to thousands of journals from around the world and from prestigious international publishers. Secondly, only the first three top editors: Reed-Elsevier, Taylor & Francis, Wiley-Blackwell (Macdonald, 2015) were taken into consideration. Thirdly, only those worldwide scientifically recognized journals have been considered in the sample, journals that have published at least 5-10 articles in relation to the analyzed syntax. Specifically, we considered for this analysis, articles with a great impact in the field and high rigor in the scientific approach.

Based on the search engines of the two databases, there were 3660 articles (in February 2018) that contained the words "innovation" on the one hand and "development" or "R&D" on the other hand in the title, keywords or abstract. The search was conducted only in these fields made available by the two platforms, as they synthesize the main elements treated in each article, representative of a written content. A search on the text would have called into question the relevance of the found articles, as the sought syntax and words could have even appeared in the article reference section, their use being considered contextual or peripheral, unrelated to the subject. Starting from the outlined above principles, we reduced the initial number of articles to 335 articles.

Table 1. Selected editors and journals					
No crt.	Editor	Journal	Abbreviation	Papers	
	Reed- Elsevier	Agricultural Systems	AS	5	
		Energy Policy	EP	20	
		Futures	Ftr	14	
1		Journal of Business Research	JBR	11	
		Journal of Cleaner Production	JCP	11	
1		Research Policy	RP	51	
		Technological Forecasting and Social Change	TFSC	23	
		Technovation	TeVa	43	
		World Development	WD	13	
	Taylor & Francis	European planning studies	EPS	24	
		Innovation management policy & practice	IMPP	17	
2		Regional Studies	RS	10	
		Technology Analysis & Strategic Management	TASM	25	
3	Wiley- Blackwell	Business strategy and the environment	BSE	6	
		Creativity and innovation management	CIM	5	
		Journal of Product Innovation Management	JPIM	25	
		R&D Management	RDM	32	
4	Total	-		335	

Source: authors

2.2 Creating the dictionary

It is one of the most sensitive steps, involving a particular workload. Given the extremely large area of the two words in question ("innovation" and "development") and the large number of articles (335 articles, spread across a period of 40 years), we aimed to define through the dictionary, areas for both "innovation" and "development". To illustrate the large volume of analysis, we present below within Table 2, the occurrence frequency of the first 7 terms within the 335 articles.

No. crt.	Term	Frequency	Cases	% Cases
1.	Innovation	28476	334	99.70%
2.	R&D	17607	292	87.16%
3.	Development	13124	332	99.10%
4.	Research	11584	334	99.70%
5.	Technology	10692	318	94.93%
6.	Firms	10420	286	85:37%
7.	Knowledge	8499	306	85.37%

Source: authors

DOI: 10.24818/18423264/55.1.21.01

8

Taken alone, the 28476 occurrences of the term "innovation" say nothing or very little, especially since there is a very broad use of the term in a variety of contexts and with multiple meanings. Thus, in the establishment of the innovation and development dictionary, two or more terms (one of which is "innovation" or "development") have been taken into account, so as to narrow the coverage of the term. The longer the syntax length, the more it has a higher degree of specificity. Thus, 36076 syntaxes have been identified which cover a broad spectrum of the "innovation" or "development" terms. Subsequently, these syntaxes were structured into 40 categories (shown in Table 3), then in subcategories, depending on the coverage.

No	Catagory	No	Catagory	
crt.		crt.	Calegory	
1	Accessibility	21	Local_&_Regional_&_Global	
2	Analysis	22	Management	
3	Areas	23	Market	
4	Business	24	Measurement	
5	Centers_&_Platforms	25	Model_&_Pattern	
6	Changes_&_Trends	26	Organization	
7	Cooperation	27	Performance	
8	Costs_&_Efforts	28	Potential	
9	Description	29	Practice	
10	Design	30	Process	
11	Development_&_Innovation	31	Program	
12	Economy	32	Project	
13	Effectiveness	33	Research	
14	Environment	34	Role	
15	Features_&_Tools	35	Services_&_Products_&_Production	
16	Geographical	36	Social	
17	Government_&_Public_Institutions	37	System	
18	Industry	38	Technology	
19	Investment	39	Theory	
20	Knowledge	40	Types	

 Table 3. Categories included in the dictionary for innovation and development

Source: authors

2.3 Text mining analysis over time intervals

Once the dictionary has been established, text mining analysis is carried out for 1979 - 2000, 2000 - 2009 and 2010 - 2018 periods of time. For each time interval, the following aspects are discussed:

• Frequencies of occurrence. The absolute frequency and relative frequency of occurrence for the dictionary subcategories and also the number of articles in which they appear.

- Clusters and dendograms. Innovation and development syntaxes can be grouped into second order clusters, which are formed as a consequence of the fact that these syntaxes were in the same context or in similar contexts. In order to identify the relations among syntaxes, Jaccard's coefficient of similarity is used (Almeida Medeiros et al., 2016). The graphical representation of these clusters appears as a dendrogram.
- Centroids and ties. It implies presenting the most significant cluster of the time period and analyzing the relationships between the dictionary categories using a network graph. It will include the subgroups within the most significant cluster, joined together by branches. Furthermore, each network graph will provide information on the existence of one or more centroids, triadic configurations or unilateral ties.

Following the steps presented in the methodology, we can analyze the results below.

3. Results

We chose to analyze the results obtained over 3 time periods, 1979-1999, 2000-2009 and 2010-2018. Although they are not equal in number of years nor in the number of offered articles (56 articles in the first period, 86 articles in the second period, 193 articles in the third period), they point to the moments when the interest in the innovations and development suffered a change, either by decreasing the number of articles (from 1999 to 2000) or by suddenly increasing their number (2009 to 2010).

3.1 Analysis for 1979 – 1999 period of time

a) Frequencies of occurrence. From Table 4, we note that the topics of interest in the field of innovation and development were based at that time on the study of innovative production, products and services, innovative processes and technologies, as well as on the management elements in the field: plans, policies and strategies.

No	Category	Frequency	% of	Number of
crt.	Category	requeitcy	shown	cases
1	Services_&_Products_&_Production	567	14.46%	39
2	Process	460	11.73%	46
3	Technology	341	8.7%	45
4	Management	308	7.86%	39
5	Types	295	7.53%	44

 Table 4. Frequency of top 5 categories of phrases (1979 – 1999 period)

Source: authors

With a similar frequency to the Management category, the articles also addressed different types of development and innovation, while costs, namely

development and innovation efforts, accounted for 6.4% of the total works of that period.

b) Clusters and dendograms. The dendogram obtained from the analysis of articles written during the period 1979-1999, reveals the existence of nine clusters (Figure 1), formed of subgroups.



Figure 1. Dendogram of co-occurrence of categories (1979-1999 period) Source: authors with Wordstat

For example, the first cluster is one consisting of two elements: Local_Regional_&_Global_I and Management_I. The cluster has a Jaccard index of approximately 0.4, indicating a low co-occurrence of syntaxes from those two categories. The second cluster consists of 11 subgroups with a general Jaccard index of 0.4. The index is almost 0.8 for the Costs_&_Efforts R&D and Features_&_Tools_R&D subgroup. Therefore, there is a high frequency of occurrence of these two syntaxes in the same context, and the groups of articles comprising them have a high degree of similarity. This subgroup joins the term Process_R&D, and beyond, the newly formed subgroup can also encounter in a particular context, the Types_R&D category. As the subgroups merge, the Jaccard Index begins to decrease, and the frequency of co-occurrence also diminishes.

c) Centroids and ties. Figure 2 presents the second cluster formed out of 11 subgroups. In this graph, each category is plotted in relation to each other based on the relative co-occurrence given by the Jaccard coefficient. Within this cluster there are two centroids (by centroid here we understand a node with the highest number of connections to neighbor nodes), "Process_I" and "Process_R&D". So the strongest

connections appear around these two categories from the dictionary. On the left hand side of the graph, "Costs_&_Efforts_R&D" and "Features_&_Tools_R&D" have a Jaccard Index of 0.778, so there is a high co-occurrence frequency for the syntaxes of these two categories in the same context. The articles containing the syntaxes have a high level of similarity.

On the right hand side of the graph, "Process_I" and "Technology_I" have a Jaccard Index of 0.667, revealing the strongest connection formed by this centroid.



Figure 2. Relevant cluster of the studied period – Subgroups and relationships among them

Source: authors with Wordstat

The link between the two subgroups is formed by one node containing the "Management_R&D" category, a node characterized by betweeness centrality. 3.2 Analysis for 2000 - 2009 period of time

a) Frequencies of occurrence. Table 5 reveals that the topics of interest in the field of innovation and development were based at that time on the study of innovative processes and systems, innovative technologies, services and products, as well as management elements in the field: plans, policies and strategies. Also, with regard to the development part, the categories that concern different types of development, as well as its geographical spread, are found in more than 40 articles of the 2000-2009 period.

No crt.	Category	Frequency	% of shown	Number of cases
1	Process_I	440	5.4 5%	59
2	System_I	388	4.8 0%	42
3	Types_D	381	4.7 2%	49
4	Local_&_Regional_&_Global_I	346	4.2 8%	40
5	Local_&_Regional_&_Global_D	336	4.1 6%	40

Table 5. Frequency of top 5 categories of phrases (2000 – 2009 period)

Source: authors

DOI: 10.24818/18423264/55.1.21.01

12

With a frequency of about 4%, the works also addressed aspects of costs, namely development and innovation efforts, which were covered in 43 works of the total works of the period.

b) Clusters and dendograms, The dendogram obtained from the analysis of articles published in the 2000-2009 period reveals the existence of 7 clusters (Figure 3). One can observe that the fourth cluster has a subgroup of Technology and Types (both related to innovation) with an index of close to 0.7. Therefore, there is a high frequency of occurrence of these two syntaxes in the same context, and the groups of articles containing them have a high degree of similarity.



Figure 3. Dendogram of co-occurrence of categories (2000-2009 period) Source: authors with Wordstat

This subgroup joins the term Process_I, so the newly formed subgroup can encounter both Management_I and System_I in a context. As the subgroups merge, both the Jaccard and the frequency of co-occurrence begin to decrease.

c) Centroids and ties. Figure 4 includes the subgroups within the fourth cluster connected by branches on which the values of the Jaccard coefficient appear.



Figure 4. Relevant cluster of the studied period Source: authors with Wordstat

Within this cluster there is one centroid, Technology_I, since this category forms six relationships with neighboring nodes. The cluster also has many triads. For example, Technology_I forms five closed triads with Description_I, System_I, Process_I, Management_I and Types_D. These triadic configurations reveal possible topics of discussion in the studied articles, often forming the basis for defining the theoretical part in terms of innovation. The strongest link is formed between Technology_I and Types_I, according to the Jaccard index of 0.7. Thus, we can appreciate a high level of co-occurrence in the same context for the elements from those two categories.

3.3 Analysis for 2010 – 2018 period

a) Frequencies of occurrence. One can observe in Table 6 the first 5 syntaxes with the highest occurrence frequency during the 2010-2018 period. The main issues of interest in the field of innovation and development focused, in the most recent period of the study, on innovative services and innovative products or processes and on management in the field of innovation. Compared with the previous periods, the absolute occurrence frequencies for the first 5 syntaxes have considerably increased, given that the number of articles in which they were found increased as well. For example, out of 193 papers published between 2010 and 2018, in 144 of them the words Process and Management appear, both related to innovation. The categories that concern different types of development, but also the development of services and products, can be found in more than 200 articles from that period.

No	Category	Frequency	% of	Number of
crt.	8- 2	1	shown	cases
1	Services_&_Products_&_Production_I	1503	6.31%	119
2	Process_I	1409	5.91%	144
3	Management_I	1196	5.02%	144
4	Cooperation_R&D	1164	4.88%	60
5	Types_D	1151	4.83%	141

 Table 6. The frequency of top 5 categories of phrases (2010 – 2018 period)

Source: authors

14

With a frequency of 4.88%, the articles also addressed issues on cooperation in research and development, these being found in 60 works in total published work at that time.

b) Clusters and dendograms. The dendogram obtained from the analysis of articles published in the period 2010-2018 (Figure 5) presents 5 clusters, composed of several subgroups. For example, the first cluster is the largest one, consisting of 12 categories and respectively 11 subgroups. Most elements included in this cluster concern innovation, as most elements of the third cluster refer to research and development.



Figure 5. Dendogram of co-occurrence of categories (2010-2018 period) Source: authors

The highest frequency of occurrence in the same context is given by the subgroup of Management_I and Process_I, therefore the articles containing them have a high degree of similarity.

c) Centroids and ties. Figure 6 shows a network graph revealing all the relationships that occur between the elements, traced as branches, each branch having a value of the Jaccard index.



Figure 6. Relevant cluster of the studied period Source: authors with Wordstat

Within this cluster, a centroid, Process_I, forms 10 surrounding links. The network graph presents a multitude of triadic configurations and a single unilateral tie between Analysis_I and Process_I, which is relatively weak considering the 0.481 value of the Jaccard coefficient. The link that indicates the greatest level of syntaxes' co-occurrence in the same context is the one that occurs between the Management_I and Process_I categories for which there is a Jaccard coefficient of 0.725.

4. Discussion

In the following, the results will be discussed according to four dimensions, seen as different aspects that can help shaping an overview on research in innovation and development.

4.1 Lexical and semantic dimension

For the three studied time intervals, we selected the most relevant cluster of the period. Both for the first timeframe, 1979-1999 and for the third one, 2010-2018, the most representative clusters are composed of 12 elements, respectively 11 subgroups. The difference between them lies in the complexity of the links among the cluster components. Thus, in the representative cluster for the 2010-2018 period, the number of triadic configurations is much higher as compared to one available for the 1979-1999 period. As we advance on the time axis, the number of triadic configurations in clusters, increased considerably. This is normal, based on the increasing complexity, interdependence and its implications on the concepts of innovation and development.

Regardless of the analyzed period, the following 4 categories of the dictionary were among the top 10 as having the highest frequency of occurrence in the articles' text: Process_I, Services_&_Products_&_Production_I, Management_I, Types_I. This is explained by the fact that in the first category, Process_I, are

included the syntaxes: Process_innovation, innovation_process and innovation_activities with very high occurrence frequencies, a cumulative value of about 1000 occurrences. In the second category,

Services_&_Products_&_Production_I, the syntax "product_innovation" is included, with a very high occurrence frequency, which proves to be of particular importance both in the literature and in practice. Further, in the third category, the syntax with the highest frequency of occurrence is "innovation_management", while in Types_I category, "national_innovation" appears 359 times.

4.2 Geographic dimension

In terms of geographical distribution of articles, it was made after the affiliation of the first author of each article. Thus, the geographical distribution shows a large coverage in Europe, followed by North America and Asia, then Africa, Australia and South America.

In Europe, the top five countries with the most publications related to the area of innovation and development are: United Kingdom with 46 papers, followed by Germany with 29 articles, the Netherlands with 22 papers, Italy with 17 and Spain with 16. At the opposite pole there are countries like the Czech Republic, Greece, and Romania, each with only one paper, and Poland and Slovenia with two papers. The Netherlands, Germany and the UK are in the world's top 10 countries according to the Global Innovation Index report of 2017. The first two European countries with a very high human development level in 2015 (according to United Nations Development Programme, Human Development Report 2016), report an average number of articles covering the innovation-development area as follows: Norway with 8 papers and Switzerland with 5 papers.

Moreover, if referred to the geographical extent of the countries, one can appreciate that the number of reported works is inversely proportional to countries geographical extent, for the first three countries in the world: Russia with two papers, Canada with 12 papers, USA with 38 papers.

A group of northern countries, Denmark, Finland and Sweden, can be formed in order to obtain 35 articles whose authors have expressed concerns in this area. If one associates the data with the one provided by the European Innovation Scoreboard (European Commission, 2017), the top three countries with the largest share of articles, the UK, Germany and the Netherlands along with the Nordic group, are part of the Innovation Leaders category, while Italy and Spain, with fewer articles, are in the Moderate Innovators category.

4.3. Economic dimension

If we try to find a connection between science / research and contemporary economic reality, it is interesting to underline that the evolution of articles by years seems to be influenced by the level of expenditures made to support research and development. The evolution of the indicator measuring the Research and development expenditure (% of GDP) at a global level (according to The World

Bank Group, data from 2018), can be found in figure 7. Thus, we are able to observe its increasing trend for 1996-2001 period, from 1.97% in 1996 to 2.08% in 2001. From 2001 to 2005, the recorded values of the indicator are on a declining trend. It seems to be recovering from 2006, but it just alternates on an increasing trend from year to year, reaching 2.22% in 2014.



Figure 7. Number of articles versus R&D expenditure as a % of GDP Source: authors

Basically, an increase in research and development expenditure entails a further increase in innovation research and hence in the number of published articles. When the level of expenditure decreases from one year to another, the effect also spreads in the interest shown for the research, the number of articles decreasing as well. For example in 2007 as compared to 2006 expenditure levels dropped, so a year later the number of articles began to decline. The situation stays the same for the next few years.

4.4 Time dimension

In terms of time elapsing, one can see the timid beginnings of research in innovation, shortly before 1980, as only one article in Futures journal of author Pavitt (1979) can be found. During this period of early research in the field, Futures journal also publishes an introduction to a special issue on the Kondratiev cycles (Freeman, 1981).

Since 1995, the field has received more attention, so until 2001, authors of 61 papers bring together the concepts of innovation and development or innovation and R&D. There is a short period afterwards, when the interest in this topic is reduced, so that in the beginning of 2005 it will show up again in a series of publications, remaining on an upward trend so far. The 1979-2001 period provides 63 works, almost 18.8% of the total, while the rest of the published articles appear after 2001. Thus, the 2001-2018 period provides 81.2% of the total published articles on the innovation – development theme.

5. Conclusions

The purpose of this analysis was to provide a full view on current research activity and salient facets related to innovation and development. The reported findings reveal valuable insights to those researchers or practitioners engaging in this field in order to guide further scientific activities. In the same time, they generate a broad picture of the changing interest in the field throughout the time.

So, some interesting facts were found in each step of the analysis. For instance, the countries that have provided the largest number of articles for this analysis are the Netherlands, Germany and the UK, also included in the Very High Social Progress category (according to the Social Progress Index) and part of the Innovation Leaders category (according to European Innovation Scoreboard)(2017).

As we advance on the time axis from 1979 to 2018, the number of triadic configurations in clusters, considerably increase, as a consequence of increasing complexity, interdependence and implications on the concepts of innovation and development.

We also tried to find a connection between science or research and contemporary economic reality. It was interesting to see that the evolution of articles by years seemed to be influenced by the level of expenditures made to support research and development. Without attempting to exhaust the field of analysis, we can state that this research identified some dimensions of investigation (temporal, economic, geographical and lexical - semantic) for which suggestive connections were presented. These dimensions, by their nature, cover a broad spectrum of human activity, generating an almost complete radiography of the "innovationdevelopment" relationship research map.

As for the limitations of this study, we admit the fact that by analysing articles indexed only in Scopus and Web of Science, we may have omitted relevant research as reference to be included in the research agenda. In the same time, we are fine with the notoriety of these two databases for offering influential literature in any research field. In the end, this analysis is expected to boost more innovative ideas when it comes to bibliographic material study and to scholarly communication flow to be explored.

ACKNOWLEDGMENTS:

This paper was co-financed by The Bucharest University of Economic Studies during the PhD program. It is also a result of the research within the 2018-2020 Postdoctoral advanced research program from The Bucharest University of Economic Studies.

REFERENCES

- [1] Aghion, P., Howitt, P. A. (1992), Model of Growth through Creative Destruction. Econometrica, 60(2): 323-351; <u>http://dx.doi.org/10.2307/2951599</u>
- [2] Almeida Medeiros, M., Barnabé, I., Albuquerque, R., Lima, R. (2016), What Does the Field of International Relations Look like in South America? Revista Brasileira de Política Internacional, 59(1): 1-31;
- [3] Aylen, J. (2010), Open versus Closed Innovation: Development of the Wide Strip Mill for Steel in the United States during the 1920s. R&D Management, 40: 67-80;
- [4] Bagchi-Sen, S., Scully, J.L. (2004), The Canadian Environment for Innovation and Business Development in the Biotechnology Industry: A Firm-Level Analysis. European Planning Studies, 12(7): 961-983. https://doi.org/10.1080/0965431042000267867;
- [5] Cernian, A., Carstoiu, D., Olteanu, A. & Sgarciu, V. (2016), Assessing the Performance of Compression Based Clustering for Text Mining. Economic Computation and Economic Cybernetics Studies and Research, 50(2): 197-210;
- [6] Cicea, C., Marinescu, C, Moroianu, N. (2015), Innovation Process and Business Functions' Implication in the New Product Development Process. In Proceedings of the International Management Conference, Faculty of Management, The Bucharest University of Economic Studies, Bucharest, Romania, 9(1): 227-233;
- [7] Cicea, C., Marinescu, C., Albu, C.F., Balan, P.D. (2019), Applying Bibliometric Mapping and Clustering on Research Regarding Biomass Related Innovation, in Proceedings of 33rd IBIMA Conference: 10-11 April 2019, Granada, Spain, pp. 2404-2419;
- [8] European Commission. (2017), European Innovation Scoreboard 2017, <u>http://ec.europa.eu/DocsRoom/documents/24829;</u>
- [9] Feng, Z., Chen, W. (2018), Environmental Regulation, Green Innovation, and Industrial Green Development: An Empirical Analysis Based on the Spatial Durbin Model. Sustainability, 10(1): 223. https://doi.org/10.3390/su10010223;
- [10] Freeman, C. (1981), Introduction. Futures, special issue: 239-245;
- [11] Fritsch, M., Franke, G. (2004), Innovation, Regional Knowledge Spillovers and R&D Cooperation. Research Policy, 33: 245–255. https://doi.org/10.1016/S0048-7333(03)00123-9;
- [12] Garfield, E. (1980), Bradford's Law and Related Statistical Patterns. Essays of an Information Scientist, 4: 476-483. http://garfield.library.upenn.edu/essays/v4p476y1979-80.pdf;

- [13] Hall, L.A., Bagchi-Sen, S. (2002), A Study of R&D, Innovation, and Business Performance in the Canadian Biotechnology Industry. Technovation, 22: 231–244. https://doi.org/10.1016/S0166-4972(01)00016-5;
- [14] Horbach, J. (2016), Empirical Determinants of Eco-Innovation in European Countries Using the Community Innovation Survey. Environmental Innovation and Societal Transitions, 19: 1-14. <u>https://doi.org/10.1016/j.eist.2015.09.005;</u>
- [15] Hudson, J., Minea, A. (2013), Innovation, Intellectual Property Rights, and Economic Development: A Unified Empirical Investigation. World Development, 46: 66–78. https://doi.org/10.1016/j.worlddev.2013.01.023;
- [16] Kozak, M.W. (2013), Innovation, Tourism and Destination Development: Dolnośląskie Case Study. European Planning Studies, 22(8): 1604-1624, https://doi.org/10.1080/09654313.2013.784597;
- [17] Macdonald, F. (2015), These Five Companies Control More Than Half of Academic Publishing. <u>https://www.sciencealert.com/these-five-companies-</u> control-more-than-half-of-academic-publishing;
- [18] Marinescu, C., Ciocoiu, C.N., Cicea, C. (2015), Drivers of Eco-innovation within Waste Electrical and Electronic Equipment Field. Theoretical and Empirical Researches in Urban Management, 10(4): 5-18;
- [19] Pavitt, K. (1979), Technical Innovation and Industrial Development 1. The New Causality. Futures, 11(6): 458-470. https://doi.org/10.1016/0016-3287(79)90044-2;
- [20] Potts, T. (2010), The Natural Advantage of Regions: Linking Sustainability, Innovation, and Regional Development in Australia. Journal of Cleaner Production, 18(8): 713-725, <u>https://doi.org/10.1016/j.jclepro.2010.01.008;</u>
- [21] Reveiu, A. & Dardala, M. (2013), The Role of Universities in Innovative Regional Clusters. Empirical Evidence from Romania. Procedia-Social and Behavioral Sciences, 93: 555-559;
- [22] Shi, Y., Tang, Y.R., Cui, L.X. & Long, W. (2018), A text Mining Based Study of Investor Sentiment and Its Influence on Stock Returns. Economic Computation and Economic Cybernetics Studies and Research, 52(1): 183-189;
- [23] Smeureanu, I. & Bucur, C. (2012), Applying Supervised Opinion Mining Techniques on Online User Reviews. Informatica economica, 16(2): 81-91;
- [24] Stankevice, I.; Jucevicius, G. (2013), Institutional vs. Sectoral Dimension of Innovation Strategies of Firms. Technological and Economic Development of Economy, 19(S360-S382);
- [25] The Global Innovation Index. (2017), 2017 Report Rankings, https://www.globalinnovationindex.org/gii-2017-report#;
- [26] Wang, C.H., Chin, Y.C., Tzeng, G.H. (2010), Mining the R&D Innovation Performance Processes for High-tech Firms Based on Rough Set Theory. Technovation, 30(7-8): 447–458. <u>https://doi.org/10.1016/j.technovation.2009.11.001.</u>