T. CIANO,
University "Mediterranea" of Reggio Calabria (Reggio Calabria, Italy; e-mail: tiziana.ciano@unirc.it)
P. FOTIA,
University "Mediterranea" of Reggio Calabria (Reggio Calabria, Italy;
e-mail: pasquale.fotia.digies@unirc.it)
B.A. PANSERA,
University "Mediterranea" of Reggio Calabria (Reggio Calabria, Italy;
e-mail: bruno.pansera@unirc.it)
M. FERRARA,
University "Mediterranea" of Reggio Calabria (Reggio Calabria, Italy; e-mail: massimiliano.
ferrara@unirc.it), Bocconi University (Milano, Italy; e-mail: massimiliano.ferrara@unibocconi.it)*

INVENTORS DYNAMICS IN BALKANIC AREA: EVIDENCES BY A NETWORK ANALYSIS

Abstract: Patent data is a key source of information for innovation economists. In recent decades it has been possible to observe its significant diffusion and success mainly thanks either to archives digitization or to authorities' greater openness with respect to patent granting procedure. Furthermore, the use of this information over time has not been limited to simple statistics on patents and their classification, but, going further, has extended to the analysis of applicants, inventors, citations, and much more. By this seminal paper, we are going to analyze starting from Data analysis related to a selection of Balkanic Countries, chosen among the most dynamic in innovation process and production of patents: Croatia, Serbia, and Bosnia and Herzegovina. How it will explain into the work, this selection was not accidental: the aim was to represent the evolution of these Countries, in terms of patent internationalization, depending on their "link" with the European Union, not all Western Balkan Countries are in fact part of it. Croatia, an official EU member since 2012, was chosen as the representative state of European influence. Some interesting results were obtained with a novel approach by social network analysis techniques.

Keywords: knowledge inventors, patents, balkan countries, social network analysis

INTRODUCTION AND LITERATURE REVIEW

For several years, the attention of the scientific community has focused on the study and analysis of bibliometric indicators; the latter are useful for relating the various research activities of the numerous universities around the world. The interest of many researchers, in fact, has focused on the data that have measured the impact that scientific collaboration networks between universities are having in the world because this type of network not only describes the academic society but has been able to trace a clear path to learn about the structure, diffusion and evolution of knowledge in an open innovation community [1, 2]. Furthermore, in the literature there are various studies on the uses of publication and analysis of citations in the evaluation of scientific activities and some of the basic statistical properties of scientific literature, in particular the asymmetry of the distributions of publications and citations, reference time frames, and various anomalies in citation patterns from one country to another. For several years, many scholars have devoted a large part of their energies

* Corresponding Author

to the development of a similar research base and infrastructure for patent bibliometry, that is, for the use of patents, and patent citations in the evaluation of technological assets. There are striking similarities between literature bibliometry and patent bibliometry, and they are both applicable to the same wide range of problems.

Narin has shown in his work that there are striking similarities between the literature and patent distributions of national productivity, inventor productivity, reference cycles, the impact of citations and citation preferences within the country [3] (see also [12] for patient bibliometrics).

Following the idea that research can be improved using public or private financial investments, Gao et al. studied the evaluation of research in China with respect to public spending [4]. Therefore, government funding is a fundamental resource for scientific research and has made a concrete contribution to the scientific and technological development of the world. But these funds come from common taxpayers, so we need to evaluate the effectiveness of these funds. In general, policymakers use the peer review method for making assessments. Making up for peer review shortcomings, the authors propose the benchmarking assessment method, mainly guiding scientometrics indicators, for evaluating publications results and research grants use. One of the topics that have aroused our interest concerns academic relations in the Balkan area. Numerous studies have been conducted on this line that analyze the scientific production of the countries of the western Balkan area, such as Albania, Croatia, Bulgaria, Bosnia and Herzegovina, etc. There are many articles focusing on scientific disciplines, institutions and journals from these countries. For this reason, this paper intends to give the Western Balkans the prominence it deserves by studying its research productivity using a bibliometric approach. Rabkin & Inhaber were among the first scholars to analyze the scientific interactions of Argentina, Brazil and Norway in terms of citations and references to the scientific literature taken into consideration [5]. In their work they show how these three nations heavily cite the publications of the central nations as opposed to those of their own country.

Another interesting and systematic study on the development of the problem of science and technology policy in the peripheric areas of the Third World is that of Moravcsik in [11].

Starting from this research, Pravdic et al. studied the academic report of a peripheral country such as (former) Yugoslavia using three different factors: possibilities and limits of the evaluation of scientific activity; problem of the form and dimensions of science as a human activity in general; specificity of communication systems in science [6]. Crescenzi et al. in their work examine the characteristics of collaborations between inventors in the United Kingdom (UK) by observing which types of proximity − geographic, organizational, cognitive, social and cultural − ethnic − among inventors are prevalent in the partnerships that ultimately has led to technological progress [7]. By studying and using a new group of British inventors, the authors provided an analysis of the associations between these "neighborhoods" and the co-patent. The results show that while collaboration within companies, research centers and universities remains crucial, external networks of inventors are a key feature of innovation teams. Furthermore, the analysis shows that external networks are highly dependent on previous social connections, but are generally not constrained by cultural or cognitive factors. Therefore, based on some findings the authors suggest that innovation policies should, rather than focus on spatial clustering, facilitate the formation of open and diverse inventor networks. Hiring inventors has long been recognized as a learning method used by innovative companies. Palomeras and Melero, in their paper state that the characteristics of the knowledge accumulated by an inventor a in their current employment determine what hiring firms can learn from him [8]. The implication is that some inventors are more likely to be hired than their peers. The authors carried out a study on the relation between the type of knowledge embodied by inventors working at IBM and their probability of moving. Relying on patent data to track the movement of inventors between companies and to characterize the type of know-how they hold, they have identified various factors of inventor mobility, such as the quality of their work; the

complementarity of their knowledge with those of other inventors; and, to a lesser extent, their experience in key areas of the firm where the firm is not a dominant player. The results obtained confirmed the role of knowledge characteristics underpinning R&D staff mobility and suggest that learning is a relevant force in the market for inventors. Knowledge networks made up of links between elements of knowledge and social networks made up of interactions between inventors both play a key role in innovation.

Brennecke and Rank, using a multilevel network approach, the authors integrate research on the two types of networks and investigate how a firm's knowledge network affects work-related interactions between its inventors [9]. To this end, they associate inventors with specific knowledge elements in the company's knowledge network and examine how this association affects the popularity and activity of inventors in a job-related consulting network. The analysis was conducted on 135 inventors working in a German high-tech company with information derived from the company's 1031 patents. The results obtained from multilevel exponential random graph (ERGM) models show that different dimensions of knowledge derived from the firm's knowledge network shape the transfer of advice between inventors in unique ways. Therefore, in their study they demonstrate how the structural characteristics of the firm's knowledge stock influence the interpersonal interactions between its inventors, thus influencing the intra-organizational diffusion of knowledge and the recombinant possibilities of the firm. The adoption of stricter patent laws and the composition of patent rights vary from country to country according to the level of economic development [10]. A patent is a contract between an inventor and a state; it guarantees exclusive right, granted for an invention, a product or a process that makes a new way of doing something accessible, offering a new technical solution to a problem. This is a technical-legal document in which there is a detailed technical description of the object of the patent and the related claims for protection. In this case, it must contain a summary of the previous state of the art, or the technology known at the time of filing. In each country there is a national office to which it is possible to apply

for a patent; by way of example, in Italy there is the Italian Patent and Trademark Office (UIBM), based in Rome and part of the Ministry of Economic Development. Nowadays, however, companies operate in an international context and it is necessary to have protection of the value of innovation not only at the local level; the possibility of extending one's right to other countries or directly to all their respective continent and beyond, is therefore recognized, by forwarding the request to various bodies, including the European Patent Office (EPO).

Within *Figure 1* are indicated: publication number (11), publication date (43), application number (21), application date (22), technological classifications (51), designated countries (84), organization Applicant (71), inventors (72), title of the patent (54) and brief description of the patent (57). The EPO, the European Patent Office, is an active member of the Task Force on Patent Statistics led by the Organization for Economic Cooperation and Development (OECD). Other members are the World Intellectual Property Organization (WIPO), the Japanese Patent Office (JPO), the US Patent and Trademark Office (USPTO), the Korean Intellectual Property Office (KIPO), the US National Science Foundation (NSF) and the European Commission (EC). The EC is represented by Eurostat and DG Research. At the request of the Task Force, the EPO created PATSTAT as a basic dataset for statistical analysis.

RESEARCH AIM AND MOTIVATIONS

This research work aims to understand how the internationalization strategy of the inventive activity of the Balkan countries has changed over the years. In particular, it was decided to extrapolate, from the database used, the information relating to 3 of the 13 countries of the western peninsula, such as: Croatia, Serbia and Bosnia and Herzegovina. They were chosen deliberately for the purpose of analyzing the way in which their link with EU (not all western Balkan countries are EU members) has affected their evolution in terms of patent internationalization. Croatia, an official EU member since 2012, was chosen as the representative state of European influence. In order to underline the differences, Serbia, an

(19) Europäisches Patentamt / European Patent Office / Office européen des brevets

(11) **EP 3 399 810 A1**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
07.11.2018 Bulletin 2018/45

(51) Int Cl.:
H04W 56/00 (2009.01)   *H04W 36/00* (2009.01)

(21) Application number: 17169554.7

(22) Date of filing: 04.05.2017

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB
GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO
PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME
Designated Validation States:
MA MD

(71) Applicant: Telefonaktiebolaget LM Ericsson (publ)
164 83 Stockholm (SE)

(72) Inventors:
• Lindoff, Bengt
  SE-237 35 Bjärred (SE)
• Alriksson, Peter
  SE-242 31 Hörby (SE)
• Åström, Magnus
  SE-224 80 Lund (SE)

(74) Representative: Ericsson
Patent Development
Torshamnsgatan 21-23
164 80 Stockholm (SE)

(54) WIRELESS COMMUNICATION DEVICE, NETWORK NODE, METHODS AND COMPUTER
PROGRAMS FOR AIDING FINDING OF SYNCHRONISATION SIGNALS

(57)   A method of a network node of a cellular communication system for mobility from a first radio access network, RAN, operating using a first radio access technology, RAT, to a second RAN operating using a second RAT. The first RAT has synchronisation signals with a first frequency allocation and the second RAT has synchronisation signals with a second frequency allocation in relation to a network frequency for the second RAT.

The method comprises determining allocation information about the synchronisation signals of the second RAT, and transmitting a radio resource control, RRC, message including the allocation information about the synchronisation signals of the second RAT. A method for the wireless communication device is also disclosed, as well as a network node, a wireless communication device and computer programs for them.
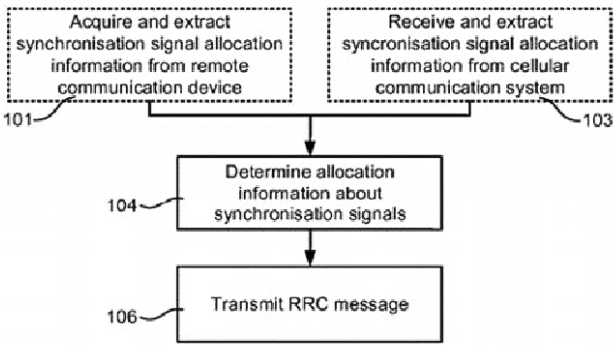
EP 3 399 810 A1

Figure 1. **Title page of a patent issued by EPO (Espacenet Patent Search – EPO)**

official EU candidate country since 2003, and Bosnia and Herzegovina, a potential candidate since 2003, were then analyzed. '' use by foreign companies, developing relationships external to the company and benefiting from an alternative mode of access to foreign markets through collaboration with other companies. The European Union knows well how important the internationalization of inventive activity is and that is why it promotes its developments, investing time and resources for the creation of incentives. The time frame chosen to evaluate the internationalization process, that is to measure the degree of

geographic heterogeneity of inventors working in the same research group, based on the country of residence (through the Country Codes), has been the one that goes from 2010 to 2017. In order to achieve this goal, an Internationalization Index of Research Groups – IGI was used, used for the first time by a group of researchers from the Complutense University of Madrid. The IGI is an innovative indicator that measures the degree of heterogeneity of the research groups. It arises from the normalization of the Herfindahl-Hirschman concentration index, used above all to measure the degree of competition present in

a given market. Homogeneity is often referred to as the opposite concept of concentration. Therefore, saying that a concentration index (such as HHI is) measures homogeneity might raise some criticism. However, it is a fact that, in this context, the concentration of the research groups, in terms of the country of residence of its members, corresponds to the homogeneity of the inventors and therefore suitable for measuring their degree. The concept opposite to homogeneity is heterogeneity. To measure the degree of heterogeneity we give a rate ranging from 0 to 1. In our case this value must be minimum (0) if each inventor of the group comes from the same country, while it must be maximum (1) if each of them is from a different country. To measure the assumed degree of heterogeneity, we first use the HHI function:

$$\sum_{i=1}^{n}(qi)^2$$

where:

n = the number of countries of residence of the inventors in the observed research group;

q_i = the share of inventors of the group redeemed in country.

For example, if we had a research group made up of three inventors, two of them Spanish and one Italian we would have:, with Spain = 2/3 = 0.66 and Italy = 1/3 = 0.33. The resulting HHI would be the sum of the squares of these two numbers (0.54). The index as structured would measure the geographic homogeneity of the inventors and would be between and, however not giving a measure in the proper sense and furthermore, groups with different n could not be compared using this index. Considering for example two research groups, where: in one we have two inventors coming respectively from Italy and China, and in the second three investors respectively from Argentina, Germany and Spain; the extent of the concentration in the two groups should be the same (i.e., the minimum), since in both each inventor is from a different country. However, HHI for the first research group is equal to 0.5, while for the second it is equal to 0.33. This is due to the fact that, as mentioned, the lower limit of the HHI is inversely related to n and the passage from one value of n to another

generates differences in the minimum value of the HHI. Moreover, given the same difference between the values of n, the difference induced in the minimum value of the HHI will make it the greater the smaller the two n are in absolute value. This means that the incompatibility problem is more serious for small values of n; to overcome this, a normalized HHI is adopted. The formula for calculating the normalized HHI is as follows:

$$H^* = \frac{N * HHI - 1}{N - 1}$$

Where is the number of inventors in the research group and HHI is the simple Herfindahl – Hirschman index calculated with the previous equation. This index ranges from 0 to 1, regardless of n. Going back to the example above, now, both hunt groups would show a equal to 0. The one described in the equation is a standardized indicator of geographical homogeneity of the research groups. To get our heterogeneity indicator – IGI, we subtract from 1 to:

$$IGI = 1 - H^*$$

The IGI measures the internationalization of a research group over a range from 0 to 1, being, as mentioned above, 0 in the case in which all inventors reside in the same country and 1 in that each of them resides in one different. Furthermore, being a standardized measure, it allows the comparison of groups with a different number of inventors. This means that, in a patent dataset, such as the one used, each patent will show its IGI score and will be comparable to all other patents based on that score. Then, once the IGI for each patent has been calculated, the result is added up and divided by the priority year, thus obtaining the average IGI for that year. In this way it is possible to describe the general trend towards internationalization of research groups in PATSTAT Global – 1992 to 2018.

EPO'S PATSTAT

The main source of patent information is the PATSTAT database. It is a database with global coverage that contains bibliographic information on almost all the patents currently in use. PATSTAT consists of two single products, such as:

PATSTAT *Global*: which contains bibliographic data relating to over 100 million patent documents from major industrialized and developing countries. It also includes legal event data from more than 40 patent authorities contained in the EPO World Legal Event Data (INPADOC), bibliographic information on applications and publications;

PATSTAT EP *Register*: contains detailed bibliographic, procedural and legal event information for EP patents (Euro-PCT published).

It is a valuable tool for the research community because it contains raw data collected transparently. This rich database promises to dramatically improve the quality of empirical research in the field.

The database consists of a set of tables following a relational database scheme in which the tables can be linked to each other using a relevant entry key. The table on patent applications, called tls201_appln, is the central element
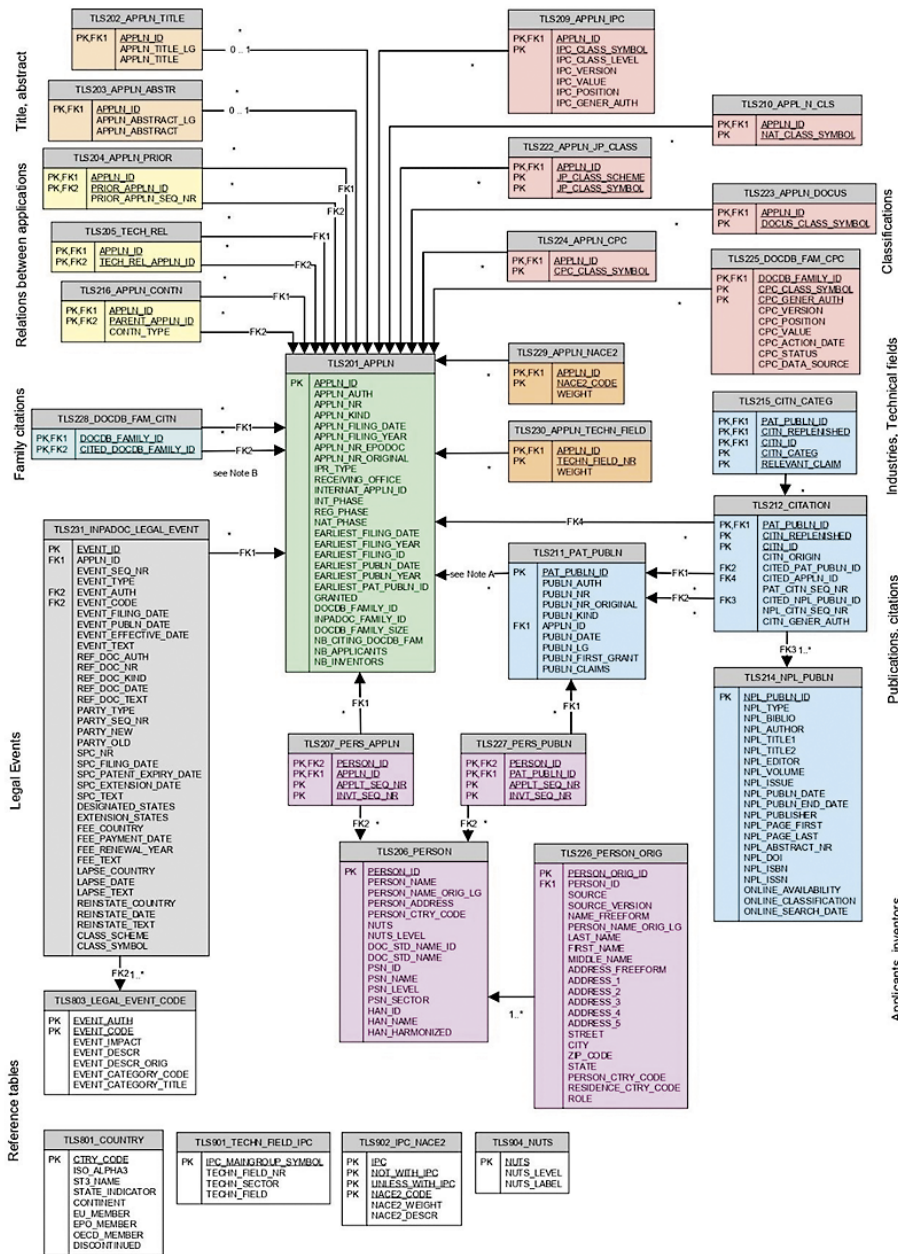

Figure 2. **Logical model diagram (Espacenet Patent Search- EPO)**

of PATSTAT. The other tables contain information on each patent application, for example, inventors and owners, technology fields, titles and abstracts, abstracts publication requests, publication requests and citations.

We see in *Figure 2* the scheme of the database.

To drastically reduce the calculation time, we run our queries on the data contained in the PATSTAT Global – Autumn Edition 2018 (hereinafter referred to as "PG light"), a database provided by EPO itself as sample data of the PATSTAT Global ranging from 1992 to 2018. PG light has the same fields as the original one but with a much smaller number of records, which allows simulations to be performed in a reasonable time, but at the same time provides less accurate data.

LIBRARIES INVOLVED IN THE WORK

Libraries are sets of written routines and functions that perform a specific task, and can be called up as needed. We can consider a library as a set of modules stored within packages. Each module contains simple instructions and definitions. The combination of various modules constitutes a library. Often the modules have already been written by other developers, and there is no need to start over each time. One of the first libraries imported into Jupyter Lab was Pandas. Once installed, the latter must be imported into the Python environment. With the code:

```
import pandas as pd
```

Pandas, and in particular its series of objects such as Series, DataFrame, which are based on the array structure, provide efficient access to the data processing activity that occupies much of a data scientist's time, providing the tools for analysis data in the Python language. The DataFrame, fundamental structure in Pandas, can be thought of as a generator of a matrix, a DataFrame can be thought of as a sequence of objects in series aligned. The package is open source and comes with different data structures that can be used for different data manipulation tasks. Pandas is a very popular library for retrieving, preparing and using future data with other libraries. It also allows you to easily retrieve data from different sources, for example: SQL database, text, CSV, Excel, JSON file. In this regard, the code used was:

```
ds= pd.read_csv(r"c:\work\patstat\tls201_appln.csv",sep=';')
df= pd.DataFrame(ds)
df.head()
```

Thus obtaining the data in tabular form of the database tls201_appln:

	appln_id	appln_auth	appln_nr	appln_kind	appln_filing_date	appln_filing_year	appln_nr_epodoc	appln_nr_original
0	18107831	ES	516894	A	1982-10-27	1982	ES19820516894	516894
1	18108793	ES	518051	A	1982-12-09	1982	ES19820518051	518051
2	18111287	ES	521001	A	1983-03-25	1983	ES19830521001	521001
3	18111288	ES	521002	A	1983-03-25	1983	ES19830521002	521002
4	18111924	ES	521787	A	1983-04-11	1983	ES19830521787	521787

Once the data is in memory, there are dozens of different operations to parse, transform, retrieve missing values, clean up the dataset, as well as SQL-like operations and a set of statistical functions to perform even a simple analysis. The next step was to import a second library: NumPy; with code:

```
import numpy as np
```

Numpy is an acronym for: Numeric Python and represents the fundamental package for scientific computing with Python. NumPy is obviously one of the largest scientific and mathematical computation libraries for Python. One of the most important features of NumPy is its array interface. This interface can be used to express images, sound waves, or other raw binary streams as arrays of real numbers with size *N*. Once the tls201_appln has been imported, we then create an array, containing all the keys of the appln_id column, and execute a search query, in the case below, all the keys associated with the HR country (Croatia).

```python
appln_id_keys= np.array(ds[c])
x= len(appln_id_keys)
x
df.colums= [column.replace(" ", "_") for column in df.columns]
HR_in_years= np.array(df.query("(appln_auth =='HR') and
(appln_filing_year < 2020)") ['appln_filing_year'])
df.query("appln_auth=='HR'")
HR_in_year
```

This series of steps made it possible to highlight the first column of the table tls201_appln, or appln_id, and to extract only patents registered in Croatia (HR) by 2020, using the appln_auth key.

	appln_id	appln_auth	appln_nr	appln_kind	appln_filing_date	appln_filing_year	appln_nr_epodoc
2953	337193977	HR	P20100151	A	2010-03-17	2010	HR2010P000151
47175	365439787	HR	P20120483	T	2012-06-08	2012	HR2012P000483T
61416	23867265	HR	P920424	A	1992-09-23	1992	HR1992P000424
61417	23867398	HR	P920702	A	1992-10-01	1992	HR1992P000702
61418	23867779	HR	P921399	A	1992-12-04	1992	HR1992P001399
...	...	...	...	...	...	...	...
234264	267454304	HR	P20080132	A	2008-03-25	2008	HR2008P000132
240300	317952050	HR	P20090047	A	2009-01-26	2009	HR2009P000047
240301	317952053	HR	P20090055	A	2009-01-29	2009	HR2009P000055
256480	334470005	HR	P20090465	A	2009-11-03	2009	HR2009P000465
268735	55557094	HR	P20070147	A	2007-04-06	2007	HR2007P000147

At this point it was necessary to use a new library in order to create a graph representing the results obtained. This is where Matplotlib comes to the rescue. This is a standard Python library used for creating 3D charts and graphs. It is quite low-level, which means that it requires more commands to generate graphs and figures than some advanced libraries. However, the main advantage is flexibility. With enough commands, you can create virtually any type of graph you want (histograms and scatter graphs to graphs with non-Cartesian coordinates). Matplotlib benefits from an additional Python library, Seabors, which enhances the data visualization tools of the Matplotlib module. The transmitted code was:

```python
import matplotlib.pyplot as plt
import seaborn
seaborn.set()
plt.title('Distribution of Croatia patents from 1992 to 2018')
plt.xlabel('years')
plt.ylabel('quantity')
x= len(HR_in_years)
x
```

By doing so, we obtained a first representative graph of all the patents made in Croatia from 1992 to 2018. The same procedure was used for other Balkan countries, such as: Serbia and Bosnia and Herzegovina; obtaining also in this case the number of patents created for that given country from 1992 to 2018.

Once these first steps have been completed, it is possible to move on to the second step of the proposed analysis. To search for inventors and their nationality for each patent, a new database table had to be imported tls207_pers_appln.

```
ds207=pd.read_csv(r"c:\WORK\PATSTAT\tls207_pers_appln.csv",sep=';
')
df207= pd.DataFrame(ds207)
df207.head()
```

From the table we note that the tls207_pers_appln has the appln_id column in common with the tls201_appln; the first column, on the other hand, person_id, contains the references of those who participated in the patent.

	person_id	appln_id	applt_seq_nr	invt_seq_nr
0	77	422617617	1	0
1	220	91957	1	0
2	220	151222	1	0
3	220	15727191	1	0
4	220	17181175	1	0

For each extracted patent we then check how many person_id are associated, and we do this by distributing the patents over some years (e.g. 2010) in order to then have results that tell me the internalization strategies of that country for that year. In Python:

```
HR_patents_year_2010= np.array(df201.query("(appln_auth =='HR') and
(appln_filing_year == 2010 )") ['appln_id'])
HR_patents_year_2010 len(HR_patent_year_2010)
person_id_HR_2010_337193977 = np.array(df207.query("(appln_id ==
337193977)")['person_id'])
person_id_HR_2010_337193977
len(person_id_HR_2010_337193977)
```

Obtained the number of person_id through the tls206_person table, we find the country code of each inventor, thus discovering their nationality and name. Therefore:

```
ds206=pd.read_csv(r"c:\WORK\PATSTAT\tls206_person.csv", sep=';')
df206= pd.DataFrame(ds206)
df206.head()
```

Getting:

	person_id	person_name	person_address	person_ctry_code	nuts	nuts_level	doc_std_name_id	doc_std_name	ps
0	1	Nokia Corporation	Keilalahdentie 4,02150 Espoo	FI	FI1B1	3	1	NOKIA CORP	2110:
1	21	Koninklijke Philips Electronics N.V.	Groenewoudseweg 1,5621 BA Eindhoven	NL	NL414	3	17	KONINKL PHILIPS ELECTRONICS NV	2270:
2	26	Lucent Technologies Inc.	600 Mountain Avenue,Murray Hill, New Jersey 07...	US	NaN	9	22	LUCENT TECHNOLOGIES INC	1779-
3	57	BAYER AG	NaN	DE	DEA24	3	51	BAYER AG	192:
4	59	HONDA GIKEN KOGYO KABUSHIKI KAISHA	1-1, Minami-Aoyama 2-chome Minato-ku,Tokyo	JP	NaN	9	52	HONDA MOTOR CO LTD	1158:

Through this series of passages and codes it was possible, for each country, to extrapolate from *database*:

```
HR_2010_inventors_337193977= np.array(df206.query("(person_id ==
42827542 or person_id == 42827543 or person_id == 42827541)")
['person_ctry_code'])
HR_2010_inventors_337193977
HR_2010_inventors_name_337193977= np.array(df206.query("(person_id
== 42827542 or person_id == 42827543 or person_id == 42827541)")
['person_name'])
HR_2010_inventors_name_337193977
```

- Number of patents;
- Number of patents per year;
- Number of inventors for each patent;
- Nationality of each inventor for each patent;
- Name of the inventor for each patent.

Thanks to this work of manipulation and research in the database, it was possible to obtain the information necessary for the calculation of the IGI.

RESULTS: SOME EMPIRICAL EVIDENCE

The first results obtained concern the quantity of patents available from 1992 to 2018 in PAT-STAT Global – Autumn Edition 2018. Out of a total of 279,881 patents, 150 were registered in Croatia, 68 in Serbia and only 9 in Bosnia and Herzegovina.

Starting from *Figure 3* on the left, it is possible to notice a significant increase in the quantity of patents produced in Croatia in 2013. It is possible to appreciate the significance of the data by contextualizing the historical events of the reference period. The entry into the EU of Croatia

(2013) in fact significantly affected the production of patents, which initially stood at relatively important values, also by virtue of an agreement signed with the EPC (European Patent Convention), but which subsequently found a significant jump in the quantities recorded. The same can be appreciated for Serbia (top right) and Bosnia and Herzegovina (bottom right). Unlike Croatia, Serbia and Bosnia and Herzegovina (*Figures 4 and 5*) have different roles towards the European Union. Indeed, Serbia registered a significant number of patents in 2012, the year in which the
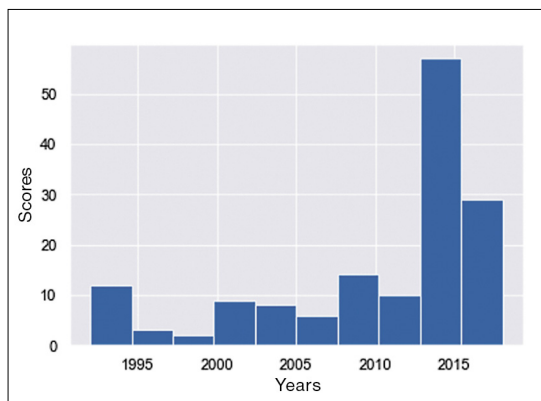

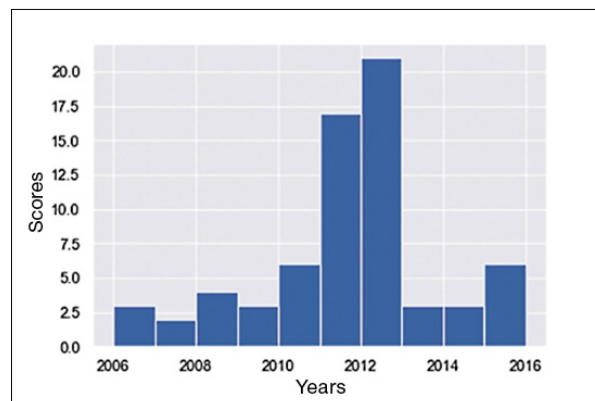*Figure 3.* **Distribution of patents from 1992 to 2018 (Croazia)**

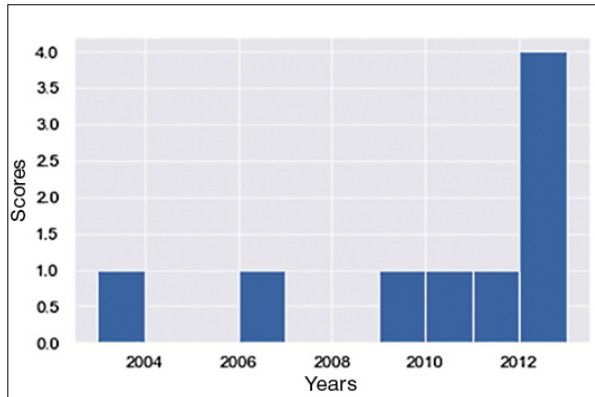*Figure 4.* **Distribution of patents from 1992 to 2018 (Serbia)**

Figure 5. **Distribution of patents from 1992 to 2018 (Bosnia and Herzegovina)**

European Council confirms Serbia as a candidate country. On the contrary, Bosnia and Herzegovina, despite having the same peak in the same year, initiated a high-level dialogue on the accession process. Therefore, it is clear that although the relations between the three countries with the EU are of a different nature, they have similarities in terms of patent production. Subsequently, for the purposes of our analysis on the internationalization of research groups, only the patents registered from 2010 to 2017 were taken into consideration and, in relation to the period of time considered, the patents registered in Croatia decreased from 150 to 100, those in Serbia from 68 to 56 and from 9 to 6 those registered in Bosnia and Herzegovina.

From *Figures 6–8* it is possible to extrapolate information about the degree of internationalization of the countries considered a few lines above. The common factor of the three countries lies in the constant growth of the IGI index, found in the years 2011–2012. As already mentioned,
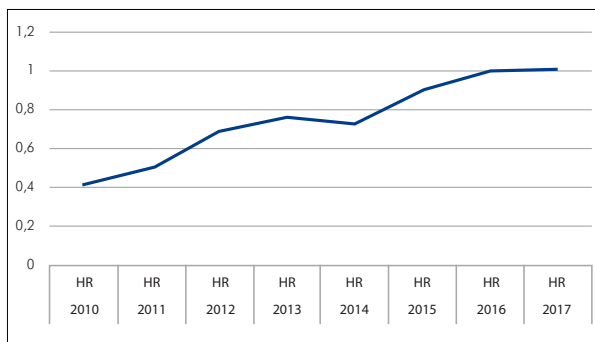
the historical context of that year is characterized by the relationship established with the EU, whose influence is visibly poured into the production of inventive activity of each country. The representative curve of the IGI index for Croatia is the only one to show constant growth over the entire time period considered; on the contrary, Serbia and Bosnia and Herzegovina present discontinuities or significant leaps in function. The IGI values for Croatia are reflected in a such a higher degree of heterogeneity that the curve's value is within a 0.4-to-1 range (the maximum degree). On the contrary, the IGI of Serbia never touches the value of 0.9, showing a significant decline after 2012. The reasons why the IGI tends to assume values close to 0 are many. One of the hypotheses would be based on a lack of collaboration with EU countries, justified by advantages not obtainable from the negotiation situation that characterized the years shown by the graph. The advantages in question could lie in market opportunities that can be achieved differently if not with facilitated partnerships between the countries of the European community. Bosnia and Herzegovina presents a completely
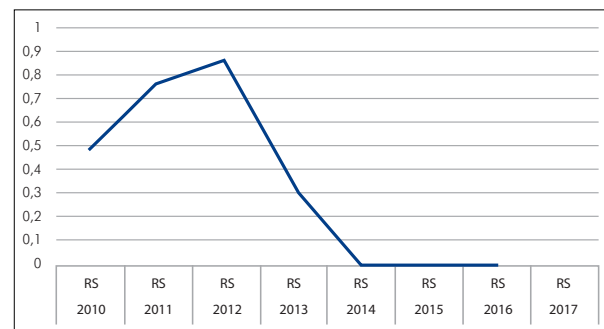

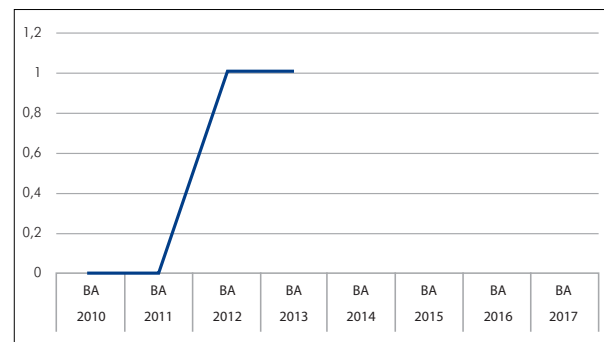*Figure 7.* **IGI per year Serbia and Herzegovin**

Figure 6. **IGI per year Croatia**

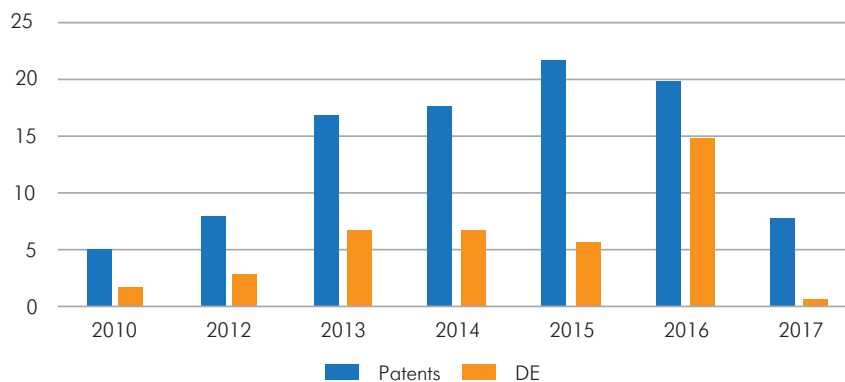

Figure 8. **IGI per year Bosnia**

Figure 9. **Percentage of German inventors out of total patents in Croatia**

different graphic situation compared to Serbia, due to the negligible number of patents available. *Figure 8* shows in fact a rapid surge of the IGI curve, allowing to reach maximum degree values in a short time. We are aware of the countless variables that could affect the different collaboration strategies between EU and non-EU countries, such that inferring a 100% reliable result is very difficult. On the other hand, it is also true that the negotiations show a degree of correlation with the internationalization index that is not negligible, making it possible to extend, although not entirely, the effect of the negotiations to the production of the inventive activity of the countries considered. Infine, tornando all'area UE si и voluto considerare il grado di collaborazione tra Germania e Croazia. *Figure 9* shows how much the levels of German influence in Croatian inventive activity are noteworthy. Although they are not characterized by a line of continuity in the proportions, settling on values which are alternately close to/lower than 50, the German influence in the patents produced in Croatia remains a constant that accompanies inventive production throughout the range temporal considered. The maximum value was recorded in the sixth year, exceeding half of the total value.

CONCLUSIONS AND FINAL REMARKS: NEW INSIGHTS FROM BIG DATA?

The analysis carried out on the PATSTAT Global database – Autumn 2018 edition made it possible to undertake a search on EU and non-EU patents, by drawing on sufficient amount of available data. The combined use of the Python computer programming language and the use of the database made it possible to grasp the potential opportunities and the predisposition of the system to provide computational support. In particular, the calculation of the IGI internalization index made it possible to numerically quantify the heterogeneity of the nationalities of inventors on European and non-European patents.

The management of the database in Jupyter enabled some cross-cutting research resulting in an extremely insightful analysis carried out in the light of an outstanding data combination mirroring the different historical backgrounds concerned. In fact, internalization is also an expression of the negotiations that take place in the community and of the internal socio-political dynamics between countries. The essence of data, or rather Big Data, contains phenomena that at first glance might not be evident; in the specific case, only the combined use of data was able to make it possible to arrive at considerations on the relationship of collaboration between the countries of the Balkan area and European countries. Although the results offer a generic picture of the degree of internalization in the Balkan area, the quality of the data available should have been at higher levels. In fact, the PATSTAT database has a level of cleanliness such that it was necessary to manage the data on a light database, thus allowing to manage the criticalities on a more modest amount of data. However, the shortage of relevant information paradoxically emphasizes the value of the results obtained which anyway can be deemed acceptable grounds for drawing reliable conclusions.

REFERENCES

1. *Eom Y.H., Fortunato S.* (2011) Characterizing and modeling citation dynamics // PloS one. 2011; 6(9).
2. *Gingras Y., Lariviere V., Macaluso B., Robitaille J.P.* (2008) The effects of aging on researchers' publication and citation patterns // PloS one. 2008; 3(12).
3. *Narin F.* (1977). Bibliometrics. Annual review of information science and technology.
4. *Gao J.P., Su C., Wang H.Y., Zhai L.H., Pan Y.*T. (2019) Research fund evaluation based on academic publication output analysis: the case of Chinese research fund evaluation // Scientometrics. 2019; 119(2): 959–972.
5. *Rabkin Y., Inhaber H.* (1979) Science on the periphery: a citation study of three less developed countries // Scientometrics. 1979; 1(3); 261–274.
6. *Pravdic N., Aganović-Boras A., Kritovac D.* (1988) In search of a "Non-citation index" indicator for scientific activity assessment in less developed countries. Case study of Croatia/Yugoslavia // Scientometrics. 1988; 14(1–2): 111–125.
7. *Crescenzi R., Nathan M., Rodriguez-Pose A.* (2016) Do inventors talk to strangers? On proximity and collaborative knowledge creation // Research Policy. 2016; 45(1): 177–194.
8. *Palomeras N., Melero E.* (2010) Markets for inventors: learning-by-hiring as a driver of mobility // Management Science. 2010; 56(5): 881–895.
9. *Brennecke J., Rank O.* (2017) The firm's knowledge network and the transfer of advice among corporate inventors – A multilevel network study // Research Policy. 2017; 46(4): 768–783.
10. *Park W.G.* (2008) International patent protection: 1960–2005 // Research policy. 2008; 37(4): 761–766.
11. *Moravcsik M. J.* (1985) Science in the developing countries: An unexplored and fruitful area for research in science studies // *4S Review*. 1985; 3(3): 2–13.
12. *Narin F.* (1994) Patent bibliometrics // Scientometrics. 1994; 30(1): 147–155.

Authors

Tiziana Ciano – Senior PhD Student, Department of Law, Economics and Human Sciences & Decision_Lab, University "Mediterranea" of Reggio Calabria; Scopus Author ID: 57215427805, ORCID: https://orcid.org/0000-0002-9854-1862 (Italy, 89124, Reggio Calabria, Cittadella Universitaria, 25; e-mail.tiziana.ciano@unirc.it).

Pasquale Fotia – Young PhD Student, Department of Law, Economics and Human Sciences & Decision_Lab, "Mediterranea" University of Reggio Calabria (Italy, 89124, Reggio Calabria, Cittadella Universitaria, 25; e-mail: pasquale.fotia.digies@unirc.it).

Bruno Antonio Pansera – Assistant Professor in Mathematical Economics, Department of Law, Economics and Human Sciences & Decision_Lab, "Mediterranea" University of Reggio Calabria; Scopus Author ID: 36830908200, ORCID: https://orcid.org/ 0000-0001-9370-7643 (Italy, 89124, Reggio Calabria, Cittadella Universitaria, 25; e-mail: bruno.pansera@unirc.it).

*Massimiliano Ferrara** – Full Professors in Mathematical Economics, Department of Law, Economics and Human Sciences & Decision_Lab, "Mediterranea" University of Reggio Calabria; Scopus Author ID: 56224779700, ORCID: 0000-0002-3663-836X (Italy, 89124, Cittadella Universitaria, 25; e-mail: massimiliano.ferrara@unirc.it); &

Full Professors in Mathematical Economics, ICRIOS – The Invernizzi Centre for Research in Innovation, Organization, Strategy and Entrepreneurship, Department of Management and Technology, Bocconi University (Italy, 20136, Milano, Via Sarfatti 26; e-mail: massimiliano.ferrara@unibocconi.it).

---

* Corresponding Author