

The Use of Complex Networks Tools to Describe the Current State of Multidisciplinary Research in Poland

Izolda Gorgol¹

¹ Department of Applied Mathematics, Lublin University of Technology, Nadbystrzycka 38D, 20-618 Lublin, Poland

e-mail: i.gorgol@pollub.pl

ABSTRACT

The purpose of this research was to specify the interrelationships of various scientific disciplines based on the official declarations of scientists. The most required scientific disciplines were identified for multidisciplinary research. Undirected graphs with the specified parameters were used for modelling of community structure. On the basis of an open database POLON, a graph was created for the connections among disciplines of science. The number of scientists in particular branches and disciplines of science was presented. The connections with in the branches of science and disciplines were shown. Four different measures were considered: percentage of double declarants, vertex degree, weighted vertex degree and betweenness centrality. The disciplines with highest level of interdisciplinarity were found, namely: biomedical engineering, informatics, culture and religious sciences, health sciences, and security sciences. On the contrary, very isolated disciplines were found, including the following: musical arts, astronomy, visual arts and conservation of cultural heritage, veterinary medicine, and archeology. Finally, it was concluded that two branches research (medical sciences or management and quality sciences) are the most interdisciplinary in Poland.

Keywords: scientific disciplines, complex networks applications, system of research, interdisciplinarity

INTRODUCTION

Complex networks are very powerful tool. Their history begins in 1959 with the papers by Erdős and Rényi [10] and Gilbert [13] which addressed the study of random networks in graph theory and were followed by Granovetter (1973), the most cited social network paper [14], the 1998 paper Watts and Strogatz on small world phenomena [35], the paper by Barabási and Albert (1999) on scale-free networks [3] and the 2003 paper by Newman on network science [24] as milestones in the field. A rapidly growing interest in complex networks is observed in the 21st century, because they appear almost everywhere. They reflect the relationships in biology (the interactions between genes, proteins, the connections between neurons), social sciences (professional, friendship ties), engineering (suppliers of energy or water grids, communication and computer networks),

economy (exchange of goods and services). They are used in predicting the spread of diseases and in preventing terrorist or cyber-attacks. One cannot forget about such huge networks as Facebook, Tweeter or LinkedIn. A fascinating description of a development of the network science is contained in a hybrid internet book by Barabási [4].

The authors used the complex networks tools to describe the current state of multidisciplinary research in Poland. The fact that according to new legal regulations, each researcher in Poland must officially declare which discipline of science he or she is working in was exploited. On 20th of July 2018, the Parliament of Poland passed a new legislation [1] concerning the whole system of high education and scientific research. It introduced many changes compared to the previous state. A short fragment of this legal background was presented. One of the legal act [30] connected with this new legislation established the classification

of branches and disciplines of science which is very close, but not totally compatible with OECD classification [26]. Currently in Poland there are 7 branches of science and 1 branch of art in Poland (marked bold). They are divided into 47 disciplines, as follows:

- **humanities**: archeology, philosophy, history, linguistics, literary studies, culture and religion sciences, art sciences,
- **engineering and technology**: architecture and urban planning, automatics, electronics and electrotechnics, information technology and telecommunications, biomedical engineering, chemical engineering, civil engineering and transport, material engineering, mechanical engineering, environmental engineering, mining and energetics,
- **medical and health sciences**: pharmaceutical sciences, medical sciences, physical cultural sciences, health sciences,
- **agricultural sciences**: forestry, agriculture and horticulture, food technology and nutrition, veterinary medicine, zootechnics and fisheries,
- **social sciences**: economics and finance, human geography and spatial management, security sciences, social communication and media sciences, political and public administration sciences, management and quality sciences, legal sciences, sociological sciences, pedagogy, canon law, psychology,
- **natural sciences**: astronomy, informatics, mathematics, biological sciences, chemical sciences, physical sciences, Earth and environmental sciences,
- **theological sciences**: theological sciences,
- **art**: film and theater arts, musical arts, visual arts and conservation of cultural heritage.

As can be seen, one additional branch/discipline is taken into account comparing to OECD, namely theological sciences. Although the branch of art is not regarded as scientific, it is taken into consideration, since it has nontrivial connections with some scientific disciplines.

Scientific disciplines are the object of study in a huge diversity of aspects. Perhaps the key issue is interdisciplinarity, itself investigated from different perspectives. Huutoniemi et al. (2009) [15] surveyed the categorizations of interdisciplinary research. For a discussion about the measures of interdisciplinarity, see e.g. Porter and Rafols (2009) [28], Porter et al. (2007) [27], Rafols and

Meyer (2009) [29], Leydesdorf (2007) [4]. Rosvall and Bergstrom (2010) [31] showed a way to observe the changes in science, including a birth of a new discipline. Kwon et al. (2017) [19] draw attention to the fact that joining the knowledge from different disciplines is not necessarily supported as much as it could be expected. Most of these studies were conducted from a global perspective. Macháček and Srholec (2019) [22] compared 174 countries and various country groups in 4 broad and 27 narrow disciplines over the period from 2005 to 2017, based on six journal-level indicators of globalization.

There are few reports about the structure of the system of scientific disciplines in particular countries. Kronneger et al. (2012) [18] described the collaboration structures in Slovenian scientific communities. Moreover, the research communities in Slovenia allowed Karlovčec and Mladenić (2015) [16] to test new interdisciplinarity measures. Local studies are mainly focused on the collaboration aspects, see e.g. Leifeld et al. (2017) [20] or Mitze and Strotebeck (2019) [23].

PROBLEM STATEMENT

The purpose of this research is a description of a narrow fragment of the whole system of the organization of science in Poland, namely interconnectedness of various scientific disciplines based on the official declarations of scientists. It is the first time in the history of research in Poland that the scholars are obliged by law to declare their disciplines of interest. The research is focused on giving a larger picture of the system, which can be a starting point for the study on its further evolution. The authors gave both quantitative and structural aspects. The scientific disciplines in Poland with the greatest potential to be involved in an interdisciplinary research were identified. At the end, the authors discussed if the official clusters of disciplines, i.e. branches of science overlap with communities of disciplines which appear during examination of a network of disciplines.

METHODOLOGY

A brief overview of the concepts used in the following part of the paper is presented below. We mainly follow van Steen (2010) [33].

General graph theory

Only undirected graphs which will be used for modelling a considered structure were described. From a mathematical point of view, a graph G is an ordered pair $G = (V, E)$, where V is a non-empty set of object called vertices and E is a family of two-element subsets of V called edges. A fact that two vertices v_1, v_2 form an edge e is denoted by a symbol $e = \{v_1, v_2\}$. In applications, vertices represent various objects, e.g. people, airports, countries and edges represent two-element relation between objects, e.g. friendship, flight connection, possessing a common border. If two vertices belong to the same edge, than they are called neighbors or it is said that they are joined by an edge. Otherwise, i.e. when a pair of vertices does not form an edge, it is said that they are disjoint. A degree of a vertex v , denoted by $d(v)$, is a number of its neighbors. A set of all neighbors of the vertex v , denoted by $N(v)$, called its neighborhood. A path is a graph with the vertex-set $V = \{v_1, v_2, \dots, v_n\}$ and the edge-set $E = \{\{v_1, v_2\}, \{v_2, v_3\}, \dots, \{v_{n-1}, v_n\}\}$. This means that the vertices are ordered and edges are present only between two consecutive vertices. The length of a path is the number of its edges. A graph is called connected if there is a path between each pair of vertices. In some applications, it is necessary to reflect the strength of the relation between two objects. For instance, let vertices correspond to airports and edges to direct flight connections among set of airports. There is an edge between two particular airports if there is at least one direct flight between them. If there is a need to represent also the number of possible flights than a weight is used. Formally, weight w is a function $w: E \rightarrow \mathbb{R}$ specified on the edge-set E with the values in a set of real numbers. In order to count all flights served by the airport (to simplify we do not pay attention to the direction of the flight) it is necessary to count not only the neighboring airports, but also the number of flights to and from each of them. For this purpose a weighted degree d_w is defined. Let $v \in V, N(v) = \{u_1, u_2, \dots, u_m\}$ and $w_i = w(\{v, u_i\})$. Then

$$d_w(v) = \sum_{i=1}^m w_i. \tag{1}$$

Note that the ordinary degree $d(v)$ can also be calculated that way by assigning the weight 1 to each edge.

Network parameters

We give here only these parameters which will be used in the following consideration. The first one is *density*. It is defined as

$$D(G) = \frac{|E(G)|}{\binom{|V(G)|}{2}}, \tag{2}$$

where: $|V(G)|$ and $|E(G)|$ denote the number of vertices and edges of the graph, respectively. Density is always between 0 and 1 and shows how close the graph G is to a complete graph, i.e. the graph in which all pairs of vertices are joined. A density close to 0 says that the graph G has rather few edges. Such a graph is called sparse. The density close to 1 says that the graph G has at most all possible edges. This graph, in turn, is called dense.

For two vertices u and v we define a *distance* between them $\text{dist}(u, v)$ as the length of the shortest path joining u and v . Now, we define a *diameter* of the graph G as $\text{diam}(G) = \max_{u,v} \text{dist}(u, v)$.

This parameter shows how *broad* a net G is and is usually considered in the context of a speed of information spreading in the net.

There are many parameters describing the role of a particular vertex in the net. One of the notion of centrality is betweenness centrality. It was independently invented by Anthonisse (1971) [2] and Freeman (1977) [12]. Let v be a vertex of a graph. The betweenness centrality of a vertex v is defined as

$$g(v) = \sum_{s \neq v \neq t} \frac{n_{st}(v)}{n_{st}} \tag{3}$$

where: n_{st} is the total number of shortest paths from a vertex s to a vertex t an $n_{st}(v)$ is the number of those paths that pass through v .

Betweenness centrality measures the number of times a vertex lies on the shortest path between other vertices. The greater this parameter is the more important a vertex is for a flow of information in the network. The basic idea is simple: if a vertex lies on many shortest paths connecting two other vertices, it is an important vertex. The reasoning is that the removal of such a vertex will directly influence the cost of the connectivity between other vertices, as other (i.e., longer) shortest paths will have to be followed (van Steen 2010 [33]).

COMMUNITIES

A network is said to have community structure if the vertices of the network can be easily grouped into disjoint sets of vertices such that there are dense connections between the vertices within sets, but sparse connections between vertices in different sets. A *modularity*, proposed by Newman and Girvan (2004) [25], is a net parameter which measures the tendency of a net to split into such sets. It is defined as the fraction of the edges that fall within the given sets minus the expected fraction if edges were distributed at random.

The community structure detection algorithms try to find dense subgraphs in directed or undirected graphs, by optimizing some criteria, and usually using heuristics. There are a large number of such algorithms. Some of them are implemented in R environment by Csardi and Nepusz (2006) [8]. See, e.g. Yang et al. (2016) [36] or Stoltenberg et al. (2019) [34] for a comparison of those. The algorithms present different approaches: maximizing the modularity parameter, using random walks or cutting a graph with the edges of maximal edge-betweenness.

Findings – A network of disciplines

A Description of a Dataset

According to [1] each person employed in a public scientific institution in Poland is obliged to declare at most two disciplines in which he or she is conducting research. For the first time, these declarations were collected to the database of Ministry of Science and Higher Education of Poland at the end of 2018. This database, i.e. POLON [9] was open to public in middle of the year 2019. The information about the numbers of scientists declaring their research in particular branches and disciplines can be established from this database (see fig. 1,2). For our purpose, the information about all researchers who declared interdisciplinary research, i.e. made a declaration in two disciplines, were downloaded. They are called *double declarants*. This sub-database contains 14052 records and was downloaded on 18th of October 2019.

Basic description

Using the data contained in our database, a graph reflecting the connections among disciplines

of science was created. Namely, disciplines are vertices of the graph and two vertices are adjacent if and only if there is at least one person who declared research in both disciplines. Moreover, weight was assigned to each edge, which is the number of persons declaring research in both disciplines. The graph has 47 vertices and 512 edges with weights varying from 1 up to 1452. These 512 edges mean that almost half of $\binom{47}{2} = 1080$ possible pairs of disciplines is being worked in. The smallest degree is 3 (canon law) and the largest is 38 (management and quality sciences). The smallest weighted degree is 8 (astronomy) and the largest is 2510 (medical sciences). The whole graph is connected which means that for any two disciplines a group of potential collaborators can be found joining these two disciplines. The diameter of the graph is equal to 3. This means that in any case, three persons are sufficient in this group. For instance, for the disciplines of astronomy, agriculture and horticulture, which are disjoint, there is a path of length 3 joining them through physical sciences and philosophy. This means that there is at least one person who works in astronomy and physical sciences, at least one who works in physical sciences and philosophy and at least one working in philosophy and agriculture and horticulture. The density of the whole net is equal 0.474, which is quite high and indicates a great possibility of cooperation. A picture of the whole graph can be seen in Figure 3. It is easy to notice there that among all the connections there are two which are stronger than others, namely {medical sciences, health sciences} with the maximal weight 1452 and {economics and finance, management and quality sciences} with the weight 969. Let us remember that this means that 1452 scientists in Poland declared that they work both in medical sciences and health sciences, and 969 of them declared that they conduct their research in both economics and finance and management as well as quality sciences. Such large numbers can mean that these two pairs contain disciplines which are very close to each other. Especially, if we compare these numbers with the third weight equal to 393 between, again, medical sciences and pharmaceutical sciences. Some stronger connections can also be observed between literary studies and culture and religious sciences, biological and medical sciences, as well as mechanical and material engineering. Another phenomenon which can be observed is that the edges of the great weight do not form any

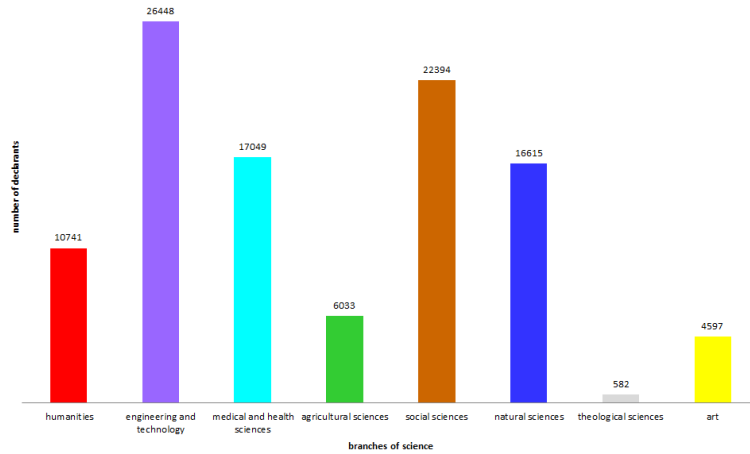


Figure 1. A number of scientist researching in particular branches of science

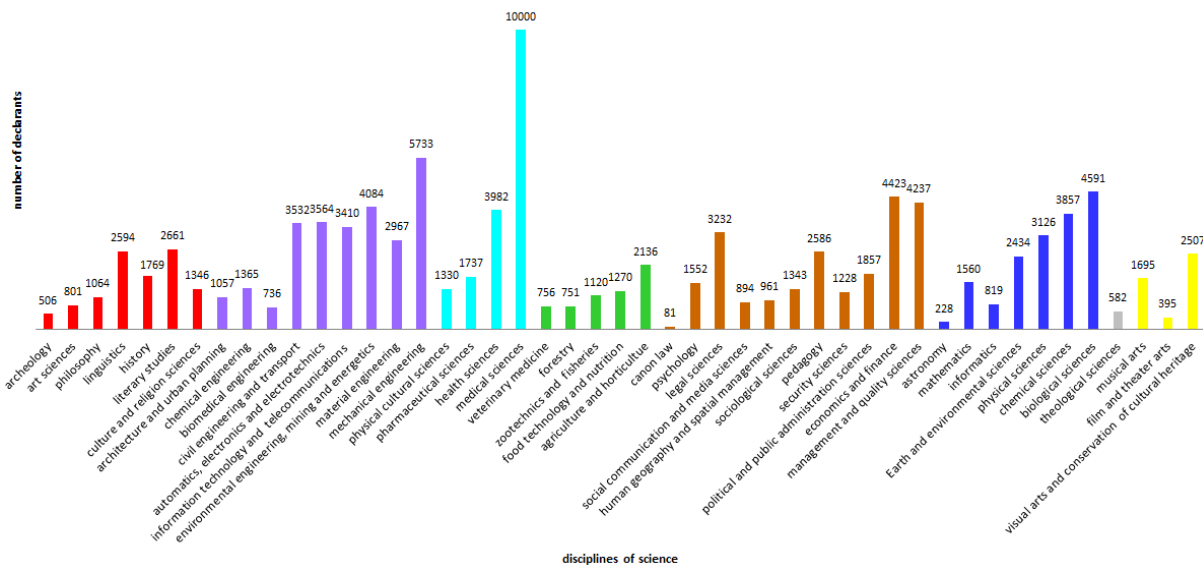


Figure 2. A number of scientist researching in particular disciplines of science. Colors are assigned to branches of science: humanities – red, engineering and technology – purple, medical and health sciences – cyan, agricultural sciences – green, social sciences – brown, natural sciences – blue, theological sciences – grey and art – yellow

triangle, which means that no three disciplines in Poland are extremely close to each other.

Inner and Intra Connections

Branches of science subsume the disciplines which are close to each other in terms of a subject and the methodology of their study. Figure 3 depicts the connections within the branches of science. Only the yellow subgraph respective to branch of art is complete, i.e. has all possible edges. There is nobody who works both in pharmaceutical sciences as well as in physical and cultural sciences within the branch of medical and health sciences. In the branch of agricultural sciences, the veterinary medicine is quite isolated,

similarly like canon law within social sciences and astronomy in natural sciences.

Another view on weight of edges

In most nets, the weight of an edge is the number of connections joining two vertices. For instance for the net of actors it is the number of films in which two actors played together, for co-authorship network it is the number of common publications, etc.

Therefore, the first choice for our net of disciplines for the weight of the edge was the number of double declarants. Of course, it shows the strength of the connection between two disciplines, but it is biased by the fact that the vertices

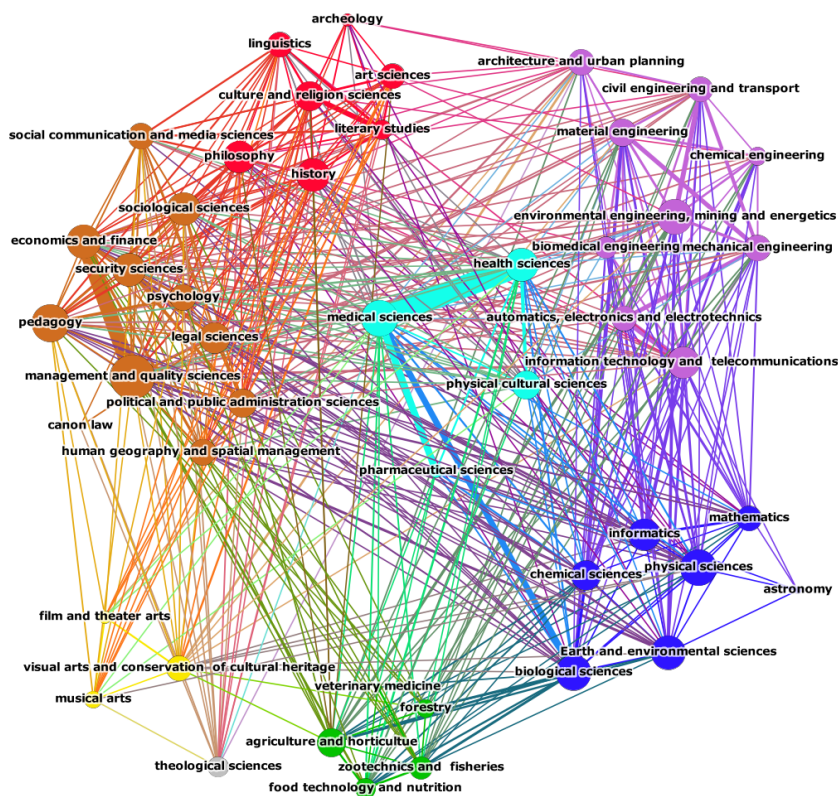


Figure 3. Connections among disciplines of science in Poland. The size of the vertex is proportional to the degree and the width of the edge is proportional to its weight. Colors are assigned to branches of science: humanities – red, engineering and technology – purple, medical and health sciences – cyan, agricultural sciences – green, social sciences – brown, natural sciences – blue, theological sciences – grey and art – yellow

are of various capacity in contrast to nets of actors or co-authorship mentioned before, where a vertex consisted of one person. In this light, some strong connections between disciplines might not be visible because of the smaller number of scholars conduct their studies in these fields. It does not mean that these disciplines are less important or useful; they can simply be not so popular at this moment. Everybody knows that there is always some kind of fashion, even in science. In order to decrease the influence of capacity of vertices and extract the hidden connections between disciplines we propose using another weight for edges, namely the Jaccard index. For two non-empty finite sets X, Y , the *Jaccard index* is defined as the ratio of the size of the intersection $X \cap Y$ and the size of the union $X \cup Y$,

$$J(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} \quad (4)$$

It is widely used as a measure of similarity between two finite sets. This approach allows showing not absolute, but relative strength of connections. The results of this approach were

presented in Figure 5. In order to underline the influence of the sizes of the sets and show the proportion to the edges in this figure, the size of the vertex corresponds to the number of declarants for each discipline and the thickness of the edge corresponds to the Jaccard index. It can be noticed that strong connections observed before remain strong, but relative strength between economic and finance and management and quality sciences is greater than that between medical and health sciences. Moreover, the first couple is more balanced, which may indicate that one could establish a common discipline “management and finance” taking advantages from both disciplines in even degree. On the other hand, the connection between medical and health sciences is much more unilateral. If these two disciplines were to merge, it would be more likely for medical sciences to absorb health sciences than for both to join on equal terms.

Apart from these conclusions some new connections, not observed so far, appeared. These are: human geography and spatial management with Earth and environmental sciences and

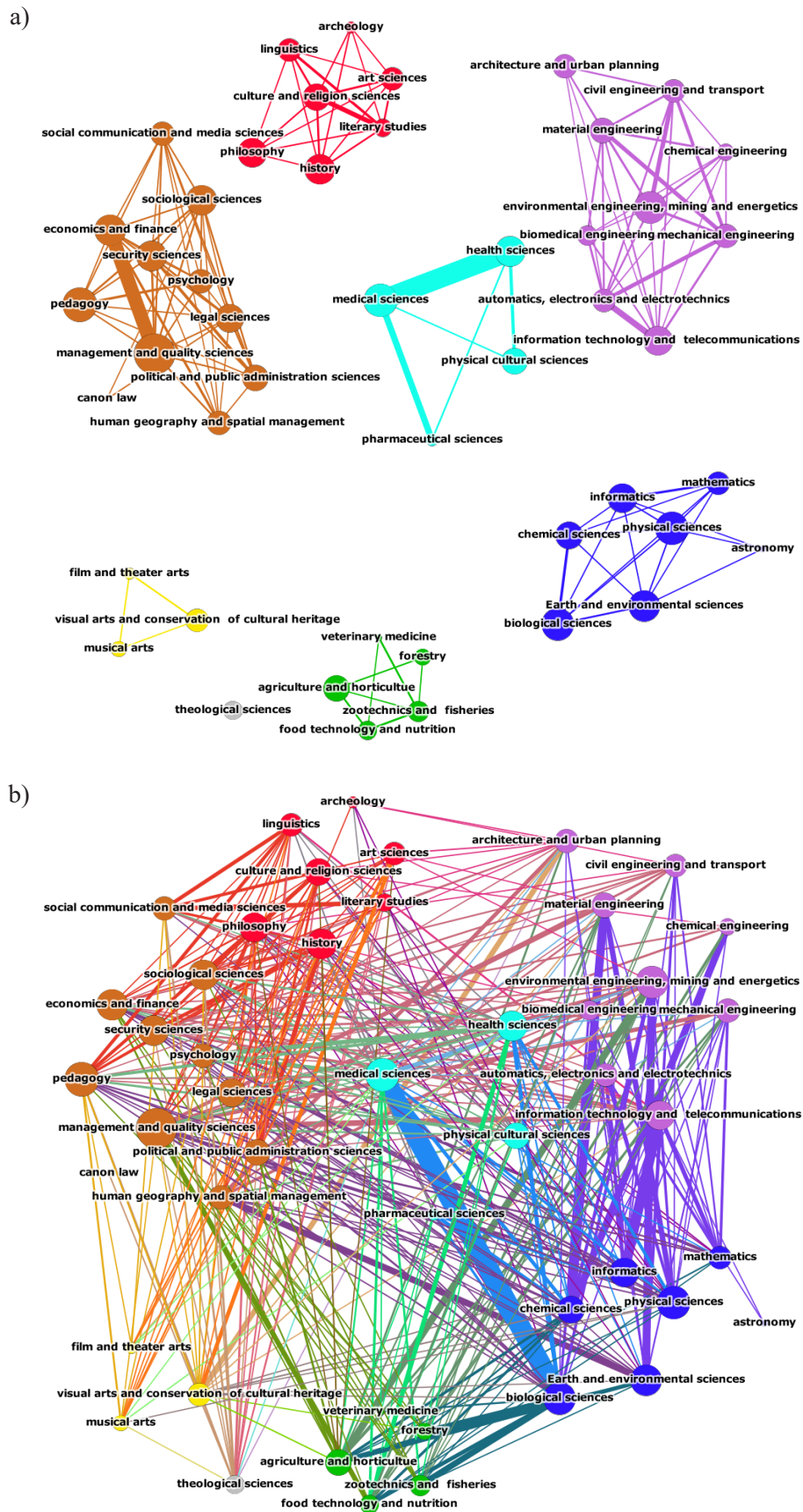


Figure 4. Connections among disciplines of science in Poland:
a) Connections within the branches; b) Connections among the branches

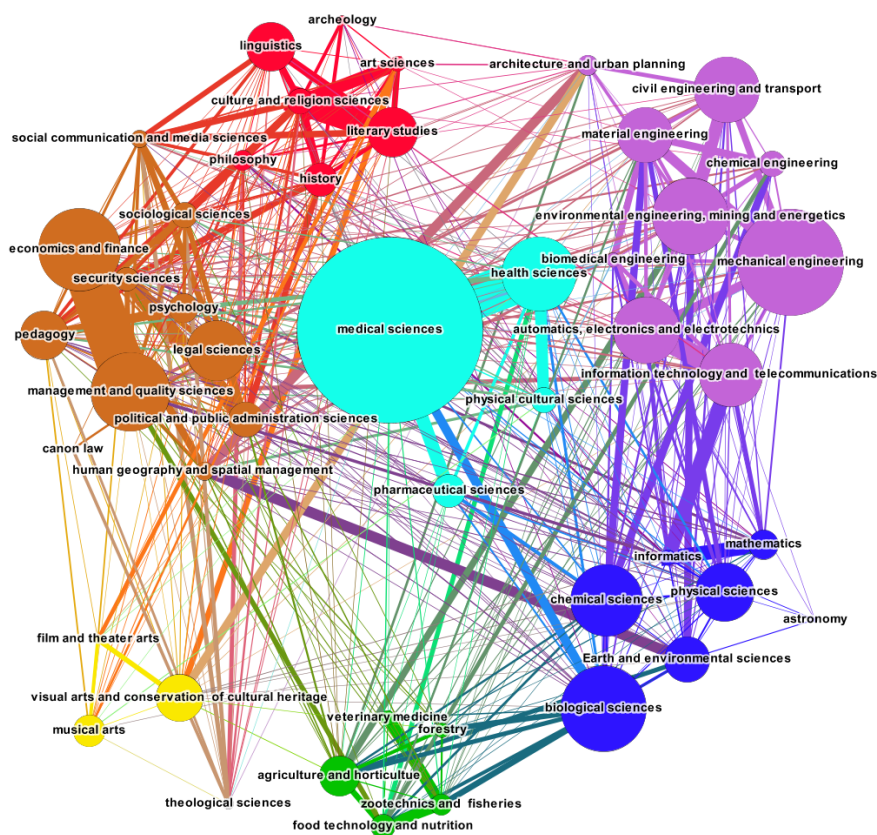


Figure 5. Connections among disciplines of science in Poland. The size of the vertex is proportional to the number of declarants and the width of the edge is proportional to its scaled weight. Colors are assigned to branches of science: humanities – red, engineering and technology – purple, medical and health sciences – cyan, agricultural sciences – green, social sciences – brown, natural sciences – blue, theological sciences – grey and art – yellow

architecture and urban planning, informatics with mathematics and information technology and telecommunications, information technology and telecommunications with automatics, electronics and electrotechnics, political and public administration sciences with security sciences and veterinary medicine with zootechnics and fisheries. These connections are not surprising and it was puzzling why they could not be seen in the previous consideration. It confirms the fact that a relative measure (the Jaccard index) is better in reflecting the strength of connection in this situation than an absolute one (the number of double declarants).

Interdisciplinarity

Defining interdisciplinarity is complex, see e.g. Barry and Born (2013) [5], Klein (2006) [17] or Siedlok and Hibbert (2014) [32]. A commonly used definition of interdisciplinarity is provided by OECD, “interaction between two or more different disciplines. The interaction may range from simple communication of ideas to the

mutual integration of organizing concepts, methodology, procedures, epistemology, terminology, data and organization of research and education in a fairly large field” (OECD, 1972 p. 25) [26]. How could interdisciplinarity be observed? Different approaches have been attempted. Perhaps the most popular one is based on the collaboration in aspects of common projects (Karlovec and Mladenec 2015 [16]) or co-authorships (Leifeld et al. 2017 [20]). The other is paying attention to bibliometrics such as references and citations in the journals which were classified in different disciplines (Porter et al. 2007 [27], Porter and Rafols 2009 [28]) or even doing text mining (Evans 2016 [11]). In this way, the interdisciplinarity of science, project, papers, etc. is described. Here, the official declarations of scholars in Poland were used. Elsewhere, we do not come across scientists describing themselves as working in one or more disciplines, rather opting out for one-discipline or multi-disciplinary labels. Moreover, such questioning was conducted for the whole scientific community. One can wonder if it reflects reality, since one person could declare at most

Table 1. Disciplines with highest and lowest percentage of double declarants:

a) highest

Discipline	Number of declarants	Percent of double declarants
Biomedical engineering	736	53.13%
Informatics	819	52.01%
Culture and religion sciences	1346	51.56%
Health sciences	3982	49.70%
Security sciences	1228	49.19%

b) lowest

discipline	number of declarants	percent of double declarants
Musical arts	1695	3.89%
Astronomy	228	4.39%
Visual arts and conservation of cultural heritage	2507	9.09%
Veterinary medicine	756	9.92%
Archeology	506	10.28%

two disciplines. On the other hand, with the high degree of specialization that we have today, perhaps only few people can conduct sophisticated research in more than two disciplines, so in our opinion this restriction is justified. Large interdisciplinary studies are performed by groups of scientists, not by single individuals. What we have in mind here, is description and measurement of the interdisciplinarity of a scientific discipline. It may raise questions, but we claim that some disciplines in Poland have a greater potential to be involved in the interdisciplinary research than others.

Percentage of double declarants

First, we will measure the level of interdisciplinarity with the percentage of scientist working in one discipline and also declaring a stake in another one. As can be seen in Table 1, according to this criterion, biomedical engineering (53.13%) is at the highest level of interdisciplinarity, followed by informatics (52.01%), culture and religious sciences (51.56%), health sciences (49.70%) and security sciences (49.19%).

At the other end of the classification there are musical arts (3.89%), astronomy (4.39%), visual arts and conservation of cultural heritage (9.09%), veterinary medicine (9.92%) and archeology (10.28%). This indicates that these disciplines are rather isolated and scholars are working mainly within just one of them.

Vertex degree

We now discuss the vertex degree as the second measure. Let us remember what the vertex

degree means for the discipline. The higher degree a particular discipline has, the greater the number of other disciplines in which at least one person works in both. Table 2 summarizes the results. The vertex degree of management and quality sciences is 38, while the maximal possible degree is 46. This means that there are only 8 disciplines that have no scientists who share their research with management and quality sciences. These are: archeology, literary studies, film and theater arts, architecture and urban planning, pharmaceutical sciences, veterinary medicine and canon law. The next disciplines with high vertex degrees are pedagogy (32), medicine sciences (31), physical sciences (31), environmental engineering, mining and energetics

Table 2. Disciplines with highest and lowest vertex degrees

a) highest

Discipline	Degree
Management and quality sciences	38
Pedagogy	32
Medical sciences	31
Physical sciences	31
Environmental engineering, mining and energetics	30

b) lowest

discipline	degree
Canon law	3
Veterinary medicine	5
Astronomy	5
Pharmaceutical sciences	10
Film and theater arts	10
Archeology	10

(30), biological sciences (30). These disciplines have something in common with at least 65% of the remaining ones. Therefore, we can claim that they are the most interdisciplinary. On the other hand, the most isolated are canon law (3), astronomy (5), veterinary medicine (5), archeology (10), film and theater arts (10) and pharmaceutical sciences (10).

The vertex degree does not reflect the strength of the connection. We claim that it can be considered as a measure of a “potential of interdisciplinarity”. When we look closer, it occurs that 108 edges of the graph have the weight 1, i.e. exactly one person has chosen two disciplines corresponding to the ends of the edge (see fig. 6). We say that this edge has a potential, it is possible to conduct research in both disciplines, but it is not popular so far.

Colors are assigned to branches of science: humanities – red, engineering and technology – purple, medical and health sciences – cyan, agricultural sciences – green, social sciences – brown, natural sciences – blue, theological sciences – grey and art – yellow

Weighted vertex degree

The weighted vertex degree does not have the disadvantage described in the previous subsection. As it is a sum of weights of a particular vertex, it takes into account the number of people for each discipline who also work in some other one. Table 3 presents disciplines with highest and lowest weighted degrees. Let us compare the results with Table 2. As for the highest ones, it can be observed that management and quality sciences and medicine sciences exchange their positions, environmental engineering, mining and energetics remains at the fifth place, but instead of pedagogy, physical sciences and biology sciences, we have health sciences, economy and finance and mechanical engineering, respectively. In each case, the replacing discipline has more declarants than the previous one (cf. fig. 2), e.g. pedagogy was declared by 2586 scholars while health sciences by 3982. However, it must be remembered, that the greater number of declarants is not a sufficient condition for a greater weighted degree. It is enough to compare health sciences (3982 declarants; second position) and mechanical

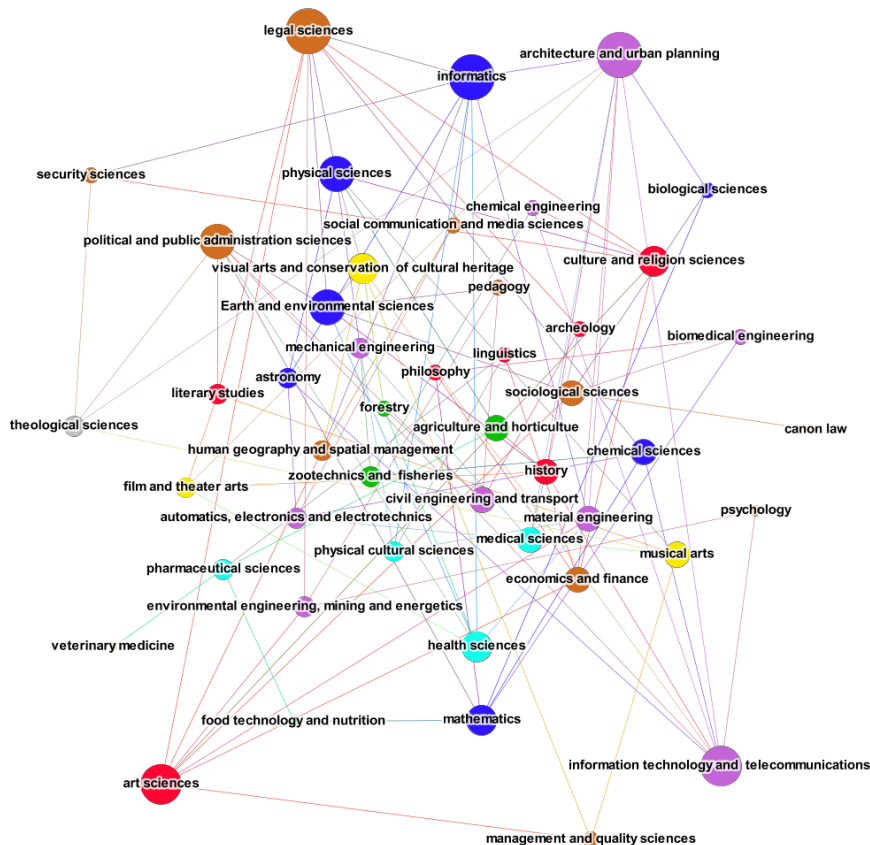


Figure 6. Connections of weight 1 among disciplines of science in Poland. The size of the vertex is proportional to the degree in this graph.

Table 3. Disciplines with highest and lowest weighted degrees

a) highest

Discipline	Weighted degree
Medical sciences	2510
Health sciences	2002
Management and quality sciences	1736
Economics and finance	1364
Environmental engineering, mining and energetics	1040
Mechanical engineering	1039

b) lowest

Discipline	Weighted degree
Astronomy	8
Canon law	22
Archeology	51
Musical arts	67
Veterinary medicine	75
Film and theater arts	77

engineering (5733 declarants, fifth position). It is also a reflection of the degree (28 and 22, respectively) and the percentage of double declarants (47.90% vs. 26.32%).

On the other hand, let us observe that the set of vertices with the smallest weighted degrees does not differ too much from the set of vertices with the smallest non-weighted degree. Both sets have five disciplines in common (canon law, astronomy, veterinary medicine, archeology and film and theater arts).

Betweenness centrality

As it is described in Subsection 2, the betweenness centrality is a measure of the importance of the vertex in the net from the point of view of the flow of information. It says what influence on that flow the deletion of the vertex would have. This parameter was considered as a measure of interdisciplinarity e.g. by Leydesdorff (2007) [4]. Tables 4(a) and 4(b) present the highest and the lowest betweenness centralities in our net of disciplines.

While looking at the parameters summarized in Table 4(a) the two disciplines appear as before: medical sciences and management and quality sciences. The next two were already seen in Table 2(a): physical and biological sciences. The last one (sociological sciences) is new in summary tables considered so far. This shows that while

Table 4. Disciplines with highest and lowest betweenness centrality

a) highest

Discipline	Betweenness centrality
Medical sciences	45.5165
Physical sciences	36.5652
Management and quality sciences	34.6785
Biological sciences	33.5191
Sociological sciences	29.6080

b) lowest

Discipline	Betweenness centrality
Astronomy	0.0000
Canon law	0.0667
Veterinary medicine	0.1429
Film and theater arts	1.4636
Archeology	1.9427

sociological sciences are not at the top of interdisciplinarity or strength of connections with other disciplines, they play an important role in the flow of scientific information and limiting research in this area would bring harm for the whole system.

Short summary of interdisciplinarity

In this section the interdisciplinarity of scientific disciplines in Poland was examined. Four different measures were considered: percentage of double declarants, vertex degree, weighted vertex degree and betweenness centrality. According to last three, connected with graph theory and reflecting the structure of the whole system, it can be concluded that two of them – medical sciences and management and quality sciences – are the most interdisciplinary, i.e. research in them can be (or must be) conducted in parallel to many others disciplines. At the other end, it can be found in all tables that astronomy, archeology and veterinary medicine are the most isolated and self-focused disciplines. Three tables also indicate canon law as such a discipline. From the point of view of the Polish system, the case of astronomy is especially interested. During the process of projecting the division of disciplines, it was proposed not to separate astronomy as an independent discipline, but treat it as a part of physical sciences. This sparked a lively discussion in the scientific world in Poland and in the end astronomy was established as a separate discipline. The research presented here confirms the correctness of this decision.

One comment should be added regarding Table 1. The disciplines occurring there can hardly be found in other tables of highest scores (except health sciences). It can be explained by joining Table 1 with Figure 5. Most of these disciplines are relatively new and can be considered as derivative or accompanying other disciplines with longer traditions (cf. health and medical sciences, informatics and mathematics, culture and religion sciences and literary studies). An isolated case is that of biomedical engineering. Its degree value is 18, only 736 researchers declare this discipline, 53.13% of whom are double declarants. It may lead to the conclusion that biological engineering in Poland is at a very early stage of development.

Communities

Finally, we compare a partition of the set of all disciplines into branches with communities obtained by means of the net-mining methods. As it was mentioned in the introduction, officially, there are 7 branches of science (humanities, engineering and technology, medical and health sciences, agricultural sciences, social sciences, natural sciences and theological sciences) and 1 branch of art in Poland. It is a traditional classification, with one exception of theological sciences. This traditional division follows from the common subject and methodology of study and, as it can be seen in Figure 3, the disciplines within branches are not necessarily strongly connected.

Therefore, we extract the clusters of disciplines which are closer to each other than to the remaining ones. We tested all algorithms provided by R environment in *igraph* package, except for the edge-betweenness algorithm, since it treats the edge weight as a distance, not as a measure of the strength of the connection. Each of them detected smaller than original $8=7+1$ number of clusters, mostly 4 or 5. The modularity of the net varied from 0.456 to 0.507, with one exception 0.018 for spin glass algorithm. Finally, as suggested by Yang et al. (2016) in [36], to establish communities we used the so-called Louvain algorithm proposed by Blondel et al. (2008) [7]. Since Gephi, an open source software for graph and network analysis [6], provides a randomized version of the algorithm, it was used to perform 50 rounds. It must be mentioned here that the one cluster appeared consisting of two disciplines – economics and finance and management and quality sciences, sometimes enriched with human geography

and spatial management. We do not present this partition because of very non-uniform distributions disciplines in clusters. On the other hand, this fact shows once again that merging economics and finance with management and quality sciences into one would have some scientific basis. Table 5 presents the comparison of the results for the algorithms used.

In order to make the best choice, modularity was used as an indicator. Two divisions were presented. First one is of the highest modularity 0.497 and 4 clusters of the most uniform sizes as follows. We gave the name to the clusters which, in our opinion, reflect their character.

- **Cluster 1 – Life** medical sciences, physical cultural sciences, health sciences, pharmaceutical sciences, biological sciences, forestry, food technology and nutrition, veterinary medicine, zootechnics and fisheries,
- **Cluster 2 – Soul** archeology, philosophy, history, linguistics, literary studies, culture and religion sciences, art sciences, social communication and media sciences, sociological sciences, pedagogy, psychology, theology sciences, film and theater arts, musical arts, visual arts and conservation of cultural heritage,
- **Cluster 3 – Tech** architecture and urban planning, automatics, electronics and electrotechnics, information technology and telecommunications, biomedical engineering, chemical engineering, civil engineering and transport, material engineering, mechanical engineering, environmental engineering, mining and energetics, astronomy, informatics, mathematics, chemical sciences, physical sciences, Earth and environmental sciences, agriculture and horticulture, human geography and spatial management,
- **Cluster 4 – Control** economics and finance, security sciences, political and public administration sciences, management and quality sciences, legal sciences, canon law.

Clusters 1–3 arose as a merge of two or three branches and the Cluster 4 is a part of sociological sciences. The Cluster 4 (control) is the most stable one. It appears in all performances of the randomized Louvain algorithm, apart from these which separated economics and finance and management and quality sciences. This indicates the strong connection among these disciplines. The second partition is chosen because it appeared in around 35% of rounds, i.e. the most often. The

Table 5. A comparison of results of basic community extracting algorithms. “yes+” denotes economics and finance and management and quality sciences cluster enriched with human geography and spatial management

Algorithm	Modularity	Number of clusters	Separate cluster for management and quality sciences and economic and finance
Optimal	0.507	4	yes +
Leading eigenvalue	0.456	5	no
Spinglass	0.018	4	yes +
Infomap	0.506	5	yes +
Label propagating	0.454	5	yes
Walktrap	0.462	7	no
Louvain (R)	0.499	4	yes
Louvain randomized (Gephi)	0.474 – 0.499	4/5	20% yes

modularity of this clustering is 0.485. In order to compare with the previous one the chemical sciences and engineering moved from Tech cluster to Life one, and the rest of Tech was split into two parts – High Tech and Space.

- **Cluster 1 – Life** medical sciences, physical cultural sciences, health sciences, pharmaceutical sciences, biological sciences, forestry, food technology and nutrition, veterinary medicine, zootechnics and fisheries, chemical engineering, chemical sciences,
- **Cluster 2 – Soul** archeology, philosophy, history, linguistics, literary studies, culture and religion sciences, art sciences, social communication and media sciences, sociological sciences, pedagogy, psychology, theology sciences, film and theater arts, musical arts, visual arts and conservation of cultural heritage,
- **Cluster 3a – High Tech** automatics, electronics and electrotechnics, informatics, information technology and telecommunications, biomedical engineering, mathematics, material engineering, mechanical engineering, physical sciences,
- **Cluster 3b – Space** astronomy, architecture and urban planning, civil engineering and transport, environmental engineering, mining and energetics, Earth and environmental sciences, agriculture and horticulture, human geography and spatial management,
- **Cluster 4 – Control** economics and finance, security sciences, political and public administration sciences, management and quality sciences, legal sciences, canon law.

Finally, we would like to point out that the only algorithm which suggested 7 clusters based on random walks produced a partition which almost overlaps the official branches of science.

CONCLUSIONS

The presented results show that the networks tools used were adequate to describe the interdisciplinarity of research disciplines in Poland. The most connected pairs of disciplines were pointed out and it was concluded that there is no strong triples in the system. Using various measures, two most interdisciplinary scientific branches were indicated – medical sciences or management and quality sciences. At the other end, isolated disciplines were identified: musical arts, astronomy, visual arts and conservation of cultural heritage, veterinary medicine, and archeology. It was shown that random walks algorithm produced clusters, which almost overlap the official branches of science in Poland.

REFERENCES

1. Act of 20 July 2018 – Law on Higher Education and Science, in Polish (Ustawa z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce), <http://prawo.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20180001668>.
2. Anthonisse J., The Rush in a Graph, Technical Report University of Amsterdam Mathematical Center, 1971.
3. Barabási A. L., Albert R., Emergence of scaling in random networks, *Science*, 286, 1999, 509–512.
4. Barabási A. L., *Network Science*, <http://network-sciencebook.com/>, Accessed 27.11.2019.
5. Barry A., Born G., Interdisciplinarity: reconfigurations of the social and natural sciences. In: Barry A., Born G. (Eds.), *Interdisciplinarity: Reconfigurations of the Social and Natural Sciences*. Routledge, Abingdon, 2013, 1–56.
6. Bastian M., Heymann S., Jacomy M., Gephi: an open source software for exploring and manipulating networks. *ICWSM*, 8, 2009, 361–362.

7. Blondel V. D., Guillaume J. L., Lambiotte R., Lefebvre E., Fast unfolding of communities in large networks, *Journal of Statistical Mechanics: Theory and Experiment*, 10, 2008, P1000.
8. Csardi G., Nepusz T., The igraph software package for complex network research, *Inter Journal Complex Systems*, 2006, Retrieved from <http://igraph.org>
9. Database of Polish Ministry of Science and Higher Education, <https://radon.nauka.gov.pl/api/katalog-udostepniania-danych>, Accessed 18.10.2019.
10. Erdős P., Rényi A., On random graphs, *Publ. Math. Debrecen*, 6, 1959, 290–297.
11. Evans E. D., Measuring Interdisciplinarity Using Text, *Socius: Sociological Research for a Dynamic World*, 2, 2016, 1–18.
12. Freeman L., A Set of Measures of Centrality Based on Betweenness. *Sociometry*, 40(1), 1977, 35–41.
13. Gilbert E. N., Random Graphs, *Annals of Mathematical Statistics*, 30, 1959, 1141–1144.
14. Granovetter M.S., The strength of weak ties, *American Journal of Sociology*, 78, 1973, 1360–1380.
15. Huutoniemi K., Klein J. T., Bruun H. and Hukkinen J., Analyzing interdisciplinarity: Typology and indicators. *Research Policy*, 39, 2010, 79–88. <https://doi.org/10.1016/j.respol.2009.09.011>
16. Karlovčec M., Mladenčić D., Interdisciplinarity of scientific fields and its evolution based on graph of project collaboration and co-authoring., *Scientometrics*, 102, 2015, 433–454.
17. Klein J. T., Afterword: the emergent literature on interdisciplinary and transdisciplinary research evaluation, *Res. Eval.*, 15(1), 2006, 75–80.
18. Kronegger L., Mali F., Ferligoj A., Doreian P., Collaboration structures in Slovenian scientific communities, *Scientometrics*, 90, 2012, 631–647.
19. Kwon S., Solomon G. E. A., Youtie J., Porter A. L., A measure of knowledge flow between specific fields: Implications of interdisciplinarity for impact and funding. *PLoS ONE* 12(10), 2017: e0185583. <https://doi.org/10.1371/journal.pone.0185583>
20. Leifeld P., Wankmüller S., Berger V. T. Z., Ingold K., Steiner C., Collaboration patterns in the German political science co-authorship network. *PLoS ONE* 12(4), 2017: e0174671. <https://doi.org/10.1371/journal.pone.0174671>
21. Leydesdorff L., Betweenness centrality as an indicator of the interdisciplinarity of scientific journals, *Journal of American Society for Information Science and Technology*, 58(9), 2007, 1303–1319.
22. Macháček V., Srholec M., Globalization of Science: Evidence from Authors in Academic Journals by Country of Origin. IDEA Study 6/2019. Institute for Democracy and Economic Analysis (IDEA), CERGE-EI, Prague. <http://www.globalizationofscience.com/> Accessed 11.12.2019.
23. Mitze T., Strotebeck F., Modeling interregional research collaborations in German biotechnology using industry directory data, *Data in Brief*, 22, 2019, 169–180.
24. Newman M. E. J., The Structure and Function of Complex Networks, *SIAM REVIEW*, 45(2), 2003, 167–256.
25. Newman M. E. J., Girvan M., Finding and evaluating community structure in networks, *Phys. Rev. E*, 69 (2), 2004, 1–15, 10.1103/PhysRevE.69.026113.
26. Organisation for Economic Co-Operation and Development Revised field of science and technology (FOS) classification in the Frascati manual DSTI/EAS/STP/NESTI(2006)19/FINAL Retrieved from (2007) <http://www.oecd.org/science/inno/38235147.pdf>
27. Porter A., Cohen A., Roessner J. D., Perreault M., Measuring researcher interdisciplinarity, *Scientometrics*, 72(1), 2007, 117–147.
28. Porter A., Rafols I., Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics*, 81(3), 2009, 719–745.
29. Rafols I., Meyer I., Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience. *Scientometrics*, 82(2), 2009, 263–287.
30. Regulation of the Minister of Science and Higher Education of 20 September 2018 on the fields of science and scientific disciplines, as well as artistic disciplines, in Polish (Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 20 września 2018 r. w sprawie dziedzin nauki i dyscyplin naukowych oraz dyscyplin artystycznych), <http://prawo.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20180001818>.
31. Rosvall M., Bergstrom C. T., Mapping Change in Large Networks, *PLoS ONE* 5(1), 2010: e8694. <https://doi.org/10.1371/journal.pone.0008694>
32. Siedlok F., Hibbert P., The organization of interdisciplinary research: modes, drivers and barriers, *Int. J. Manage. Rev.*, 16, 2014, 194–210.
33. van Steen M., Graph Theory and Complex Networks. An Introduction, Maarten van Steen (Publ.), 2010.
34. Stoltenberg D., Maie D., Waldherr A., Community detection in civil society online networks: Theoretical guide and empirical assessment, *Social Networks*, 59, 2019, 120–133.
35. Watts D. J., Strogatz S. H., Collective dynamics of “small-world” networks, *Nature*, 393, 1998, 440–442.
36. Yang Z., Algesheimer R., Tessone C. A., Comparative Analysis of Community Detection Algorithms on Artificial Networks, *Sci. Rep.* 6, 30750, 2016, <https://doi.org/10.1038/srep30750>