RedNIE Red Nacional de Investigadores en Economía

Agent Based Models and Simulation in Social Sciences: A bibliometric review

Emiliano Alvarez (Universidad de la República/Universidad Nacional del Sur)
Gabriel Brida (Universidad de la República)
Silvia London (Universidad Nacional del Sur)

DOCUMENTO DE TRABAJO 2020-26

Octubre de 2020

Citar como:
Alvarez, Emiliano, Gabriel Brida, y Silvia London (2020). Agent Based Models and Simulation in Social Sciences: A bibliometric review. *Documento de trabajo RedNIE, 2020-26*.

Agent Based Models and Simulation in Social Sciences: A bibliometric review

Emiliano Alvarez * Gabriel Brida† Silvia London‡

October 2020

Abstract

Since the first agent-based models (ABM), the scientific community has been interested in making not only the results of computational models understandable but also the modeling description, to facilitate their replication. The form that has been adopted to a greater extent has been the ODD (Overview, Design concepts, and Details) protocol, which provides a generic structure for its documentation. This protocol provides a way to clearly explain the procedures and interactions of the complex systems to be analyzed, with applications that have spread across different disciplines. This work will show a bibliometric review of the articles that emerged from the first publication of this protocol in 2006, analyzing the development that ABMs have had in the social sciences. A description will be made of the lines of research with the greatest activity and the links between them will be analyzed; while summarizing the countries, universities, and journals with the highest contributions.

Keywords: Social Simulation; Complex Systems; ODD protocol; Bibliometric Analysis

1 Introduction: documenting models in social sciences

In recent decades, the toolbox of social science researchers has been enriched with the possibility of simulating phenomena with Agent-Based Models (ABM). It is clear that, as mentioned by Gilbert[1] there is a consensus as to what social simulation allows us: the possibility of observing macroscopic processes as an emergent process that arises from the interactions between "simple" agents. For this reason, in the context of understanding social phenomena as the result of a complex web of interactions, a growing number of works attempt to conceptualize these issues with ABMs.

In this sense, Squazzoni (2010) states that using ABMs not only allows to work on well-established problems but also allows to work on other issues, for which it is only possible through a multidisciplinary approach. Therefore, ABMs have an impact not only in well-established disciplines but also encourages the creation of multidisciplinary new topics. As mentioned by Squazzoni et al. (2014), there are some differential aspects of the use of ABMs in social sciences, including

*Universidad de la República, Uruguay & Universidad Nacional del Sur, Argentina. ealvarez@ccee.edu.uy

†Universidad de la República, Uruguay. gbrida@ccee.edu.uy

‡Universidad Nacional del Sur, Argentina. slondon@uns.edu.ar

[1] "...computer programs offer the possibility of creating 'artificial' societies in which individuals and collective actors such as organisations could be directly represented and the effect of their interactions observed." (Gilbert, 2004)

the heterogeneity of behaviors (which includes the analysis of decision-making) and the analysis of social dynamics. Another relevant aspect is the importance of journals for the dissemination of these concepts, approach followed by Squazzoni and Casnici (2013); Hauke et al. (2017), who analyze the articles published in the Journal of Artificial Societies and Social Simulation (JASSS).

Now, why document it? It is important to note here that several authors show some problems that arise from the increasingly "complicated" models that are introduced to analyze complex systems with a higher degree of definition. There is a need to find a common language (Squazzoni, 2010) or a *lingua franca*, in terms of Garibay et al. (2019). This standard needs to communicate the different stages of the modeling process: building, describing, analyzing, evaluating, and replicating ABMs. This standard has to be flexible, but without falling into anarchy, as Richiardi et al. (2006) states.

Is necessary to document all these aspects of the models because they are not simply "equation-based models" which is a strongly established tradition in social sciences. Models are not clearly defined only with the equations, so it is necessary to communicate many other aspects. The intention is to make them understandable and complete, allowing their further replication.

Since 2006, a protocol for ABM documentation has begun to be promoted, which has become one of the most used in social sciences: the Overview, Design concepts and Details (ODD) protocol (Grimm et al., 2006). So, ODD is an acronym for the phases of the model documentation process, as shown in Figure 1. In the first instance, we have *Overview*, where the structure and processes of the model are detailed. Then, not only the model but also its implicit rationality and its *Design* principles are presented, for example, if certain aspects emerge from the system (endogenous) or if they are parameters of the model (exogenous). The *Details* include all the information necessary for the model to be replicated.

It is a protocol that does not arise from the social sciences[2] but has evolved, with several revisions (Grimm et al., 2010, 2020) and extensions (Müller et al., 2013). The purpose of this protocol seeks to respond to these failures in model communication, seeking to make their understanding and replication easier, at the same time that it seeks to ensure that the models are adequately described. A relevant aspect of this protocol: although it is suggested that it is suitable for ABM documentation and communication and can even be generalized to all types of models, as mentioned by Grimm et al. (2020), it is indifferent to the subject to be analyzed. Although it began to be used in ecology, nothing prevents a model that works on inflation, unemployment, productive efficiency, or inequality from being able to be communicated in this way.

This work aims to: i) describe the state of the art of documenting the ABMs in social simulation, ii) detect which topics are the most frequently used, iii) define well-established research lines, and iv) show the main articles within each area. In particular, this research seeks to analyze whether there are differences between research in social simulation that follows the ODD protocol and those who do not follow this protocol. An additional objective of this study is to know if the use of these protocols favors communication and multidiscipline. For future research, this study aims to provide a detailed description and useful information on the use of ABMs in simulation in social sciences. Likewise, bibliometric analysis allows to have an overview of the literature by identifying the best articles, authors and journals.

The paper is structured as follows. Section 2 presents the research methodology and data. Section 3 describes the results of the bibliometric analysis and Section 4 shows a further discussion, provide the conclusions, establish limitations of our study and proposes future research.

[2]It arises from ecology, which has a longer tradition in the use of ABMs. In Grimm et al. (2010) this issue is mentioned and the degree of penetration of this modeling in different disciplines is shown.

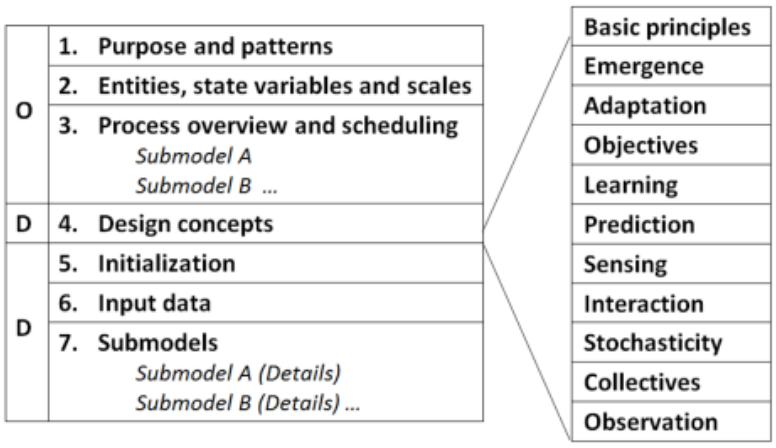

Figure 1: Structure of the model, following the ODD protocol. Source: Grimm et al. (2020).

2 Data and methodology

Bibliometric techniques are used to access quantitative information about scientific production expressed through written publications. The analysis of the bibliographic information of the articles published in journals began in the 1960s with the seminal study of Price (1965), and -as stated by Ellegaard and Wallin (2015)- the "bibliometric methods or 'analysis' are now firmly established as scientific specialties and are an integral part of research evaluation methodology especially within the scientific and applied fields"

There are different types of bibliometric analysis to be performed, based on the links between objects: bibliographic coupling analysis, co-citation analysis, co-author analysis and co-word (keywords co-occurrences) analysis. Coupling analysis establishes that two articles are linked if they have many references in common (Kessler, 1962). In the example (see Figure 2, left), we see that item A and item B refer to items C, D, E, and F, so we can assume that items A and B share information. Some authors criticize this approach, since it may also happen that different aspects of the same document are being cited, for which they may not be sharing information. Another type of analysis that we can perform from the Scopus dataset is co-citation analysis (see Figure 2, right), which states that the link between two articles depends on the number of occasions on which they are cited together (Small, 1973). Co-author analysis takes into account the number of connections that authors have within the network structure (White and McCain, 1998) while keywords co-occurrences analysis allows studying the cognitive structure of the network, as established in Callon et al. (1983), from the connections between concepts of the articles' keywords.

The data was collected from the Scopus database. It is one of the most important data bases in scientific literature, based on peer-review literature. Currently [3] It has in its records approximately 80 million articles, more than 27000 scientific journals, more than 1500 books and more than 700 conference proceedings in different disciplines. At the same time, it provides reliable and consistent metadata, with relevant information on publications, as well as about the authors and references of articles. To justify the choice of Scopus, Ahmad et al. (2020) propose some advantages of using Scopus compared to other databases. In particular, the authors state that Scopus "has very comprehensive coverage of academic journals", they sostain that "the citation coverage is also

[3] Access date: October 1st, 2020.

very comprehensive, with 1.7 billion cited references dating back to 1970" and that "the citation information from Scopus appears to be more accurate than WoS".

To perform the bibliometric analysis, articles related to ODD (Overview, Design and Concepts) were selected from Scopus Database[4]. The criterion was to seek the articles within the Scopus database in the disciplines corresponding to the social sciences that cited Grimm et al. (2006) or Grimm et al. (2010).

The data collection process had several phases. In a first phase, the total of articles that cited Grimm et al. (2006) (1527 papers) and Grimm et al. (2010) (1292 papers) were separated. From this set of papers, we consider only those related to the disciplines of Economics, Social Sciences, Business and Management and Decision Sciences. Then, after this process the number of remaining articles are 299 and 288 respectively. Finally, we eliminated duplicate articles, that is, those that cited both articles. Overall, the final sample consists of 440 articles[5]. The data collected from the Scopus database were authors, title, abstract, keywords, affiliation, country, year and references. Data analysis were performed in ***R*** (R Core Team, 2020), using *bibliometrix* package (Aria and Cuccurullo, 2017). The advantage of using this methodology is the existence of efficient statistical algorithms, the possibility of accessing different quality numerical routines and various data visualization and mapping tools. In addition, this software is an open and freeware source program.

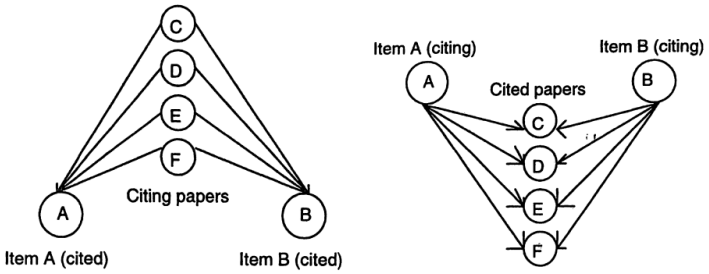

Figure 2: Examples of coupling analysis (left) and co-citation analysis (right). Source: Garfield (1988).

3 Results

The results presented in the next subsection are divided into two parts. In the first, a bibliometric analysis of the selected articles is carried out, where we carry out a quantitative description of the articles, journals, and central countries on the subject; We also carry out an analysis of the historical direct citation network and co-citation analysis, an analysis based on the keywords used and we study the collaboration networks between countries and between institutions. In the second part, we introduce a more general analysis of the ABMs in social sciences and we analyze the similarities and differences with the group of articles that follow the ODD protocol.

3.1 Summary of the selected papers

Figure 3 (left), shows that there is a steady increase in the number of annual new articles. This growth becomes more important from 2010, year when Polhill (2010); Grimm et al. (2010) were

[4]Accessed August 31, 2020.

[5]Descriptive statistics for this database can be found in the Appendix A.

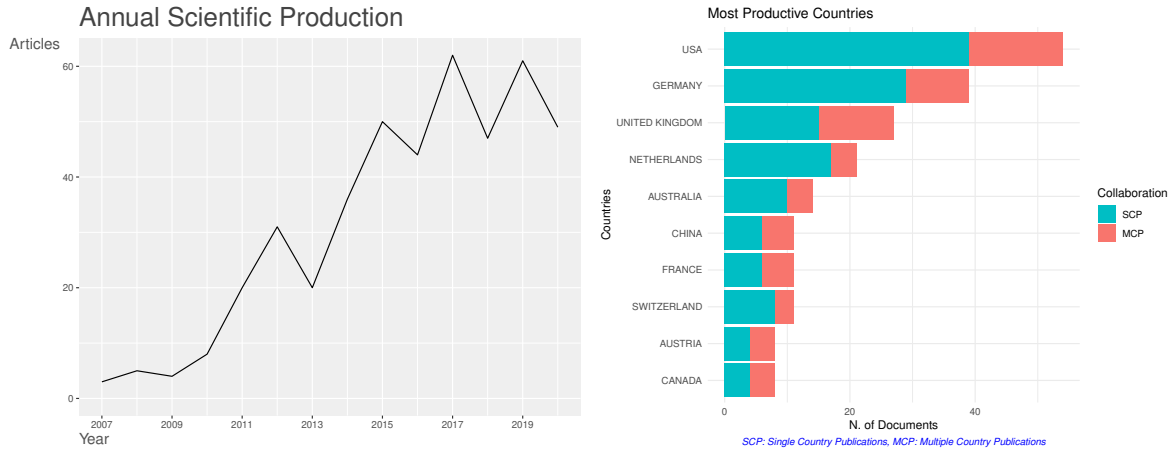

Figure 3: Annual scientific production (left) and most productive countries (right), for the analyzed articles.

published. These articles are important because they mark the first stage of the protocol updating and the publication of this update in JASSS. Figure 3 (left) shows a delay in the development of theuse of ABMs in Social Sciences, which can be attributed to various reasons, including: i) the latter development of computational Economics compared with mainstream economics disciplines, ii) the poor computational background of the economists, iii) the epistemological problem that produces the use of artificial economies. Also, we can see in Figure 3 (right) that the countries with the largest number of articles published, both in authorship and co-authorship, are from European countries and North America. In particular, USA is the most productive country, with 54 publications, and Germany is the highest producer of ABMs research in Europe. Note that Africa and the rest of America is marginal in the scientific production in this field. An interesting phenomenon is shown, which is the proportion of articles whose authors belong to more than one country. This is a result that arises from international research networks on the subject, which will be discussed below. Figure 3 (right) shows that the distribution of research concentrates on a small quantity of countries. This could be a consequence of the fact that some developed countries offer important research grants to study complex systems and thus researchers give priority to ABMs research.

An important aspect to analyze the dissemination of knowledge is its distribution within scientific publications. This will allow us to know if there are specialized journals on the topic, if these developments are concentrated only in journals specialized in a few topics or if, on the contrary, the journals where they are published contribute to dissemination. In Table 1 we observe that JASSS is the journal with the most publications in the social sciences within articles that cite Grimm et al. (2006) or Grimm et al. (2010) - considering the 203 sources reviewed, see Appendix A -, even taking into account that none of these articles were published in this journal[6]. What we can also observe is that within the ten journals with the most publications, we have journals related to simulations or issues related to ecological or environmental problems. Bradford (1934) have investigated the sources of the papers published in the field of geophysics and categorized the journals into three Zones; organizing journals in descending order and each zone has an almost equal amount of scientific papers to the field. Bradford's law establishes that the relationship follows a geometric progression: n, n^2, n^3, where n is the number of journals in Zone 1. Bradford called the

[6]Both articles were published in *Ecological Modelling*.

first zone the "nucleous" of journals devoted to the subject. In this database, the "nucleous" of the subject consists of 6 journals, zone 2 consists of 53 journals and zone 3 consists of 144 journals. It can be seen that in zone 1, 3% of the journals concentrate a third of the articles and 36.5% of citations. Zone 2 contains 26% of the journals and 37% of citations. Finally, Zone 3 contains 71% of the journals and 26.5% of citations. As can be seen in Table 1, 20% of the selected articles were published in JASSS, as well as these articles received 1218 (20.3%) of the 6006 total citations received.

Sources	Articles	Citations	Bradford classification
JASSS	88	1218	Zone 1
Landscape Ecology	16	421	Zone 1
Computers Environment and Urban Systems	14	312	Zone 1
Ecological Economics	10	97	Zone 1
Cybergeo	9	10	Zone 1
Computational and Mathematical Organization Theory	8	124	Zone 1
Journal of Cleaner Production	8	89	Zone 2
Journal of Land Use Science	8	74	Zone 2
Land Use Policy	8	69	Zone 2
Sustainability	7	25	Zone 2

Table 1: Most relevant sources, measured by the number of articles published, number of total citations and Bradford classification.

Within the keywords, we have two data series: the keywords selected by the authors and the keywords that are assigned by Scopus. We show both rankings in Table 2 and can be organized in four different categories. First, the importance of the agent-based approach is highlighted. In both classifications it is among the most used keywords, also linked with other important ones[7]. In addition, it is highlighted the importance of simulation and numerical and computational methods. Simulation is a fundamental aspect of these models. Third, note the empirical and methodological aspects, such as validation and replication. These are aspects related to the possibility of communicating and quality control of the models. Fourth, topics mentioned before appear: the analysis of decision making, economics, land use, and ecology. These results, although not surprising, also suggest that these references have not spread to areas such as finance, macroeconomic models, or opinion dynamics.

	Author Keywords (DE)	Articles	Keywords-Plus (ID)	Articles
1	agent-based model	209	computational methods	72
2	simulation	33	agent-based model	69
3	validation	13	autonomous agents	65
4	agent-based simulation	11	numerical model	49
5	individual-based model	11	decision making	40
6	social simulation	11	computer simulation	36
7	land use	9	modeling	34
8	replication	9	simulation platform	21
9	agent-based computational economics	7	land use change	19
10	climate change	7	simulation	19

Table 2: Most relevant keywords. Author Keywords (DE): author's keywords. Keywords Plus (ID): keywords associated to the manuscript by SCOPUS.

Thereupon, we generate the historical direct citation network. The aim is to represent chronologically the most relevant articles in a field of study. In particular, here we will seek to create a map and be able to observe the most relevant articles, the links between them, and which of them have helped to open new fields of study. In Figure 4 (Up) the graph is represented, generated from

[7]In this category, we also select "individual-based". In Vincenot (2018) a bibliometric analysis is carried out that shows the joint evolution of both concepts and how some articles made possible to connect different disciplines.

Article	Cluster	Betweenness centrality
Grimm et al. (2010)	1	0.14
Bonabeau (2002)	1	0
Wilensky (1999)	1	0.14
Gilbert (2008)	1	0.14
Grimm et al. (2006)	2	13.74
Macy and Willer (2002)	2	13.74
Epstein and Axtell (1996)	2	13.74
Hardin (1968)	2	9.6
Axelrod (1997)	2	13.74

Table 3: Co-citation analysis and clustering. Betweenness centrality of articles in clusters 1 and 2.

the three quartiles of articles with the highest number of citations. To represent it, we plotted it with the 150 articles with most citations. Most of the literature arises from Polhill et al. (2008). Apart from this main branch, we can see other topics that emerge, such as industrial networks (Mantese and Amaral, 2017), disease propagation (Simoes, 2012), circular economy (Ding et al., 2016), commuting patterns (Ge and Polhill, 2016), urban expansion (Li et al., 2019) and innovation and policy interventions (Arfaoui et al., 2014).

In Figure 4 (Down) we analyze only the 30 most cited articles in our database, which allows us to identify some lines of research that emerge from Polhill et al. (2008). These lines of research refer to topics such as energy efficiency (Friege et al., 2016), tourism (Johnson et al., 2017), the analysis of the results of the models, including output analysis but also calibration, validation and replication procedures (Lee et al., 2015), the feedbacks between social and natural systems (An et al., 2014), and a group that shows the advances in the methodology (Polhill, 2010).

In Figure 5 we perform the co-citation analysis previously explained, where two articles have a stronger link if they are cited at the same time by many other articles. We used Louvain's clustering algorithm (Blondel et al., 2008) and we found three clusters, which refer to three different kinds of articles: in blue (Cluster 2) we found the most cited articles of the data set and with greater centrality in the network. These articles have been the main link between ABMs and the social sciences. In red (Cluster 1) we found the articles that are the methodological references, both in the ODD protocol and in the use of ABMs as well as in the most commonly used software for modelling purposes in ABMs. In green (Cluster 3), we found a cluster of articles that are less central and that serve as a reference for land use and land change topics. In Table 3, we show the articles that belong to clusters 1 and 2 and the centrality measured from the betweenness centrality (Freeman, 1977). Vertices with high betweenness may have considerable influence within a network by their control over information passing between others. They are also the ones whose removal from the network will most disrupt communications between other vertices. From the betweenness centrality measure, it is observed that the articles belonging to cluster 2 are the central ones in this network and serve as a link between different topics and disciplines.

Now, we analyze in Figure 6 the keyword network. The links between two keywords will be stronger the more times they appear together within the data set. Using Louvain's algorithm we find four clusters, which are related to different topics: in red, we have those related to the economy and decision-making processes; in blue those related to spatial analysis; in green the keywords related to modeling and simulation, and in violet we can see computational methods and topics related with land use and ecology.

In figure 7 we can see two graphs where we can appreciate the level of regional and international collaboration of the research centers. On the left, we have the network of countries and on the right, we find the network of research centers. In the first instance, the most important research centers are in the United States and Western Europe, and almost all the research centers that work on the subject are placed there. On the other hand, the research centers that have participated in the most relevant works on the subject are in turn those who manage to connect research centers

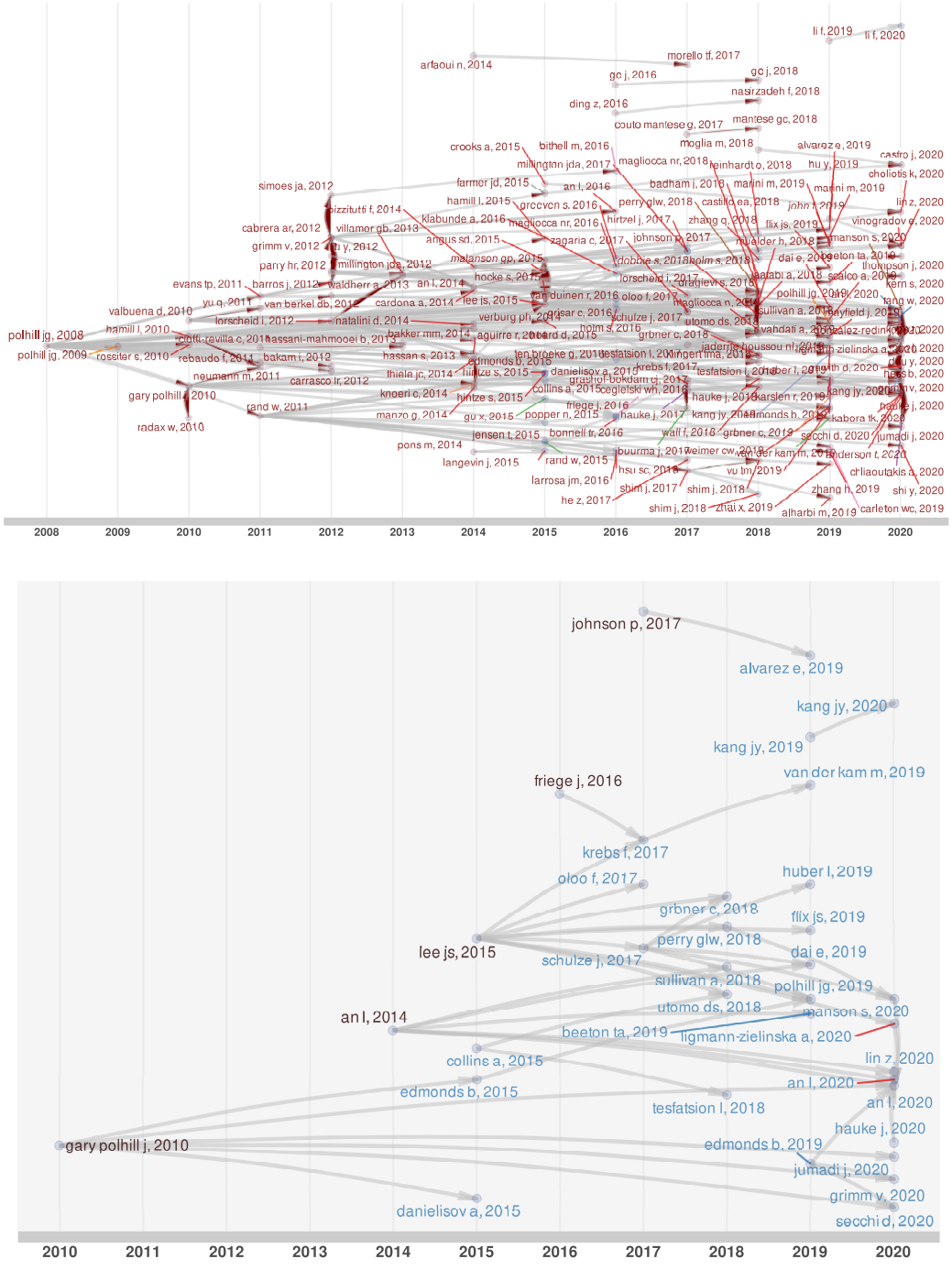

Figure 4: Historical direct citation network. Up: with 150 most cited papers. Down: with 30 most cited papers.

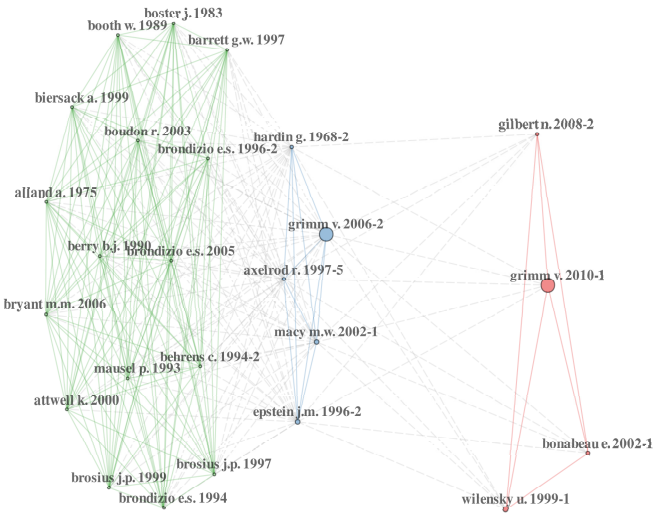

Figure 5: Co-citation analysis and clustering. Co-citation graph: cluster 1 (in red), cluster 2 (in blue), cluster 3 (in green).

from different regions of the world. In this network, we have many countries (periphery) that are linked to only one of the central countries (core) and there are many that are disconnected from the network.

3.2 ABMs in Social Sciences

In this section, we will seek to analyze the articles that use ABMs in social sciences. For this, we carry out a new search in Scopus, which includes the concepts of "agent-based" and "model", limited to articles in the same categories we mentioned above: *TITLE-ABS-KEY ("agent-based" OR "individual-based" OR "multi-agent" OR "multiagent" OR "abm*" AND "model*" OR "simulat*" AND "social*") AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (SUBJAREA , "ECON") OR LIMIT-TO (SUBJAREA , "SOCI") OR LIMIT-TO (SUBJAREA , "BUSI") OR LIMIT-TO (SUBJAREA , "DECI"))*. The asterisk (*) is used so that the search finds words with the same root but with a different ending. For example, *"model*"* finds results where the words *model*, *models*, *modeling*, *modelling* among others are used. That is especially important due to geographical differences in the use of the words *"modeling"* and *"modelling"* to refer to the same concept. That allows the search to be simpler and more understandable, as well as its subsequent analysis and replication. In total, the search allowed us to work with a data set of 1,717 articles[8].

We carried out the co-citation analysis on this dataset (see Figure 8) and found two groups, using Louvain's algorithm: in the red group, there are the most cited articles on the subject (Wilensky, 1999; Schelling, 1971; Barabási and Albert, 1999) and those that deal with the most

[8] Accessed September 10, 2020. Scopus Database.

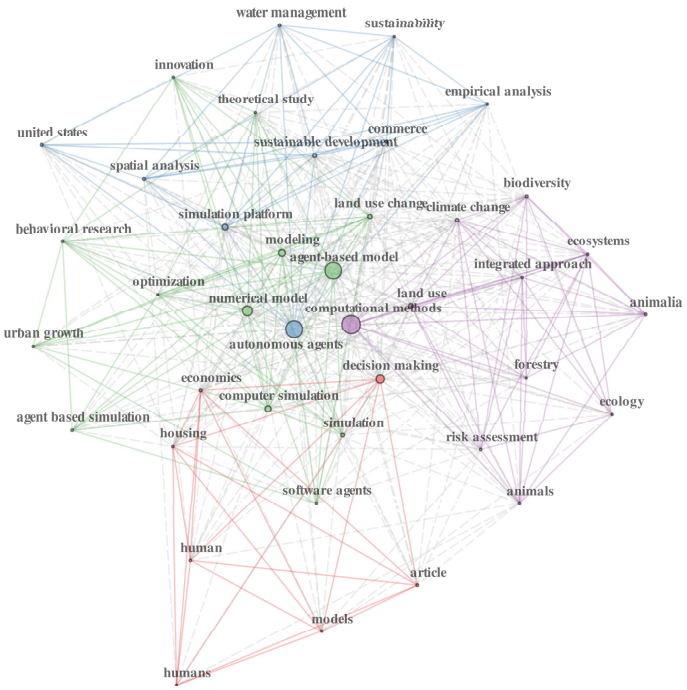

Figure 6: Keywords co-occurencies (Scopus keywords).

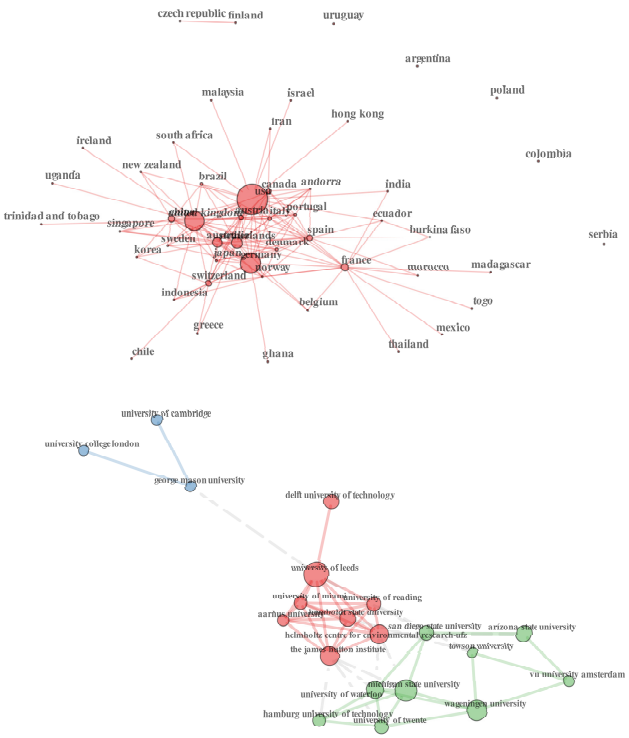

Figure 7: Collaboration networks: between countries (left) and universities (right).

important topics in ABMs in the social sciences (economics and social systems as complex systems, fundamental articles of network analysis, among others). On the other hand, the blue group is made up of some review articles and articles that seek to link the social sciences with the ABMs (Grimm et al., 2010; Farmer and Foley, 2009) and other disciplines (Parker et al., 2003; An, 2012). Here we note that Grimm et al. (2010) is weakly linked to the core of ABMs articles in social sciences. This is an indication that shows us that the ODD protocol has not yet managed to generalize its use in the social sciences, although we should also note that there are no other alternatives within the literature on ABMs in the social sciences.

When analyzing the keywords co-occurrences (see Figure 9) we will see that other words appear different from the previous data set, but that at the same time the grouping is coherent. Here we find three clusters. In red, the most used keywords are found, showing the link between ABMs and social phenomena (social influence, diffusion, social simulation, social network). In blue, we find the links to game theory (as interactive agents are represented) and complexity. In green, we found the keywords related to multidisciplinary topics (segregation, norms, learning and climate change).

What we do in the following table (Table 4) is a comparison of the most used keywords, both in the articles that follow the ODD protocol and in the articles of ABMs in the social sciences. Above we see this comparison with the author keywords while below we see the same comparison but with the words assigned by Scopus. Looking at the two classifications, we can see that there are many similarities in the most important ones (ABM, simulation, social simulation, modeling, etc). Instead, if we look at the differences, they are both in the topics covered and in the methodological emphasis: in the first case, we find land use and climate change compared to social influence, cooperation, and game theory; on the other hand, we see that validation and replication have

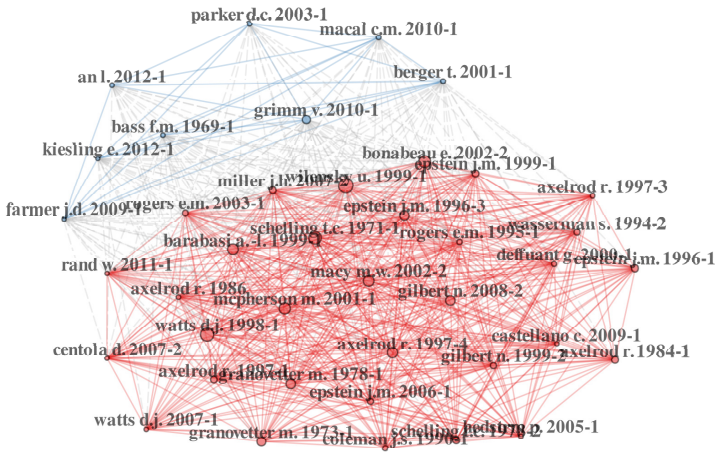

Figure 8: References co-citation analysis

greater relevance in articles that follow the ODD protocol.

Author Keywords (ODD literature)	Articles	Author Keywords (ABMs in social science)	Articles
agent-based model	209	agent-based model	646
simulation	33	social network	197
validation	13	agent-based simulation	109
agent-based simulation	11	simulation	99
individual-based model	11	social simulation	60
social simulation	11	multi-agent systems	39
land use	39	social influence	39
replication	34	complexity	34
agent-based computational economics	31	cooperation	31
climate change	28	game theory	28

Keywords-Plus (ODD literature)	Articles	Keywords-Plus (ABMs in social science)	Articles
computational methods	72	computational methods	173
agent-based model	69	agent-based modeling	160
autonomous agents	65	social network	152
numerical model	49	autonomous agents	146
decision making	40	computer simulation	146
computer simulation	36	human	136
modeling	34	decision making	123
simulation platform	21	modeling	120
land use change	19	multi agent systems	111
simulation	19	simulation	68

Table 4: Keywords' co-occurrences: comparison with articles citing Grimm et al. (2006) or Grimm et al. (2010)

In Table 5, we show the list of the 10 journals with the most publications within the ABMs dataset in social sciences. On the one hand, the journals with the most publications already appeared in the previous list, so we can say that these journals have published articles that follow the ODD protocol and articles that do not. In a way, they serve as a link between the two pieces of literature. Again, JASSS appears as the journal with the highest number of publications on the subject, which supports Squazzoni and Casnici (2013); Hauke et al. (2017) reviews. On the other hand, we find differences with the classification of the literature that follows the ODD protocol literature: now many economics journals appear (in bold), which suggests that in particular, the ODD protocol has not achieved significant advances in economics. We also appreciate that journals with a broader spectrum disappear from the list, which may mean that there are differences in the level of integration of the different disciplines. That is, we find more multidisciplinary journals and topics in the documents that cite Grimm's papers.

If we look at collaboration networks (Figure 10), as in the previous case, the most important

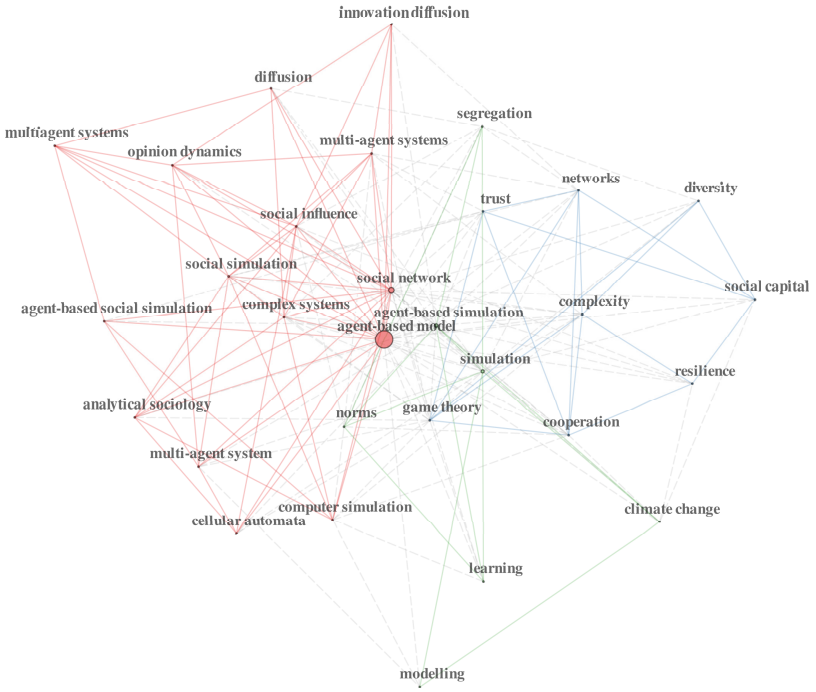

Figure 9: Author keywords.

	Sources	Articles
1	JASSS	269
2	Computational and Mathematical Organization Theory	59
3	Computers Environment and Urban Systems	30
4	Sustainability	22
5	**Social Science Computer Review**	21
6	**Computational Economics**	20
7	**Journal of Economic Interaction and Coordination**	19
8	**Journal of Economic Behavior and Organization**	18
9	Journal of Cleaner Production	17
10	**European Journal of Operational Research**	16

Table 5: Most relevant sources. In bold, sources that they do not appear among the 10 with the most publications within the selected articles.

research centers are in the United States and Europe. The country network is star-shaped (core-periphery), as in the previous case. Looking at some metrics of collaboration networks (Table 6), we can see that these are networks with low density but with high transitivity, so all of these networks are small-world networks.

On the other hand, we can observe the density of the networks. The networks of research centers and countries of the dataset that follow the ODD protocol have more links per node, have higher transitivity, and a lower average path length[9]. That is, information circulates more fluidly through this network than in the network of ABMs in social sciences, that could be explained by the presence of more specific lines of research and methodologies. Another possible explanation for this fact is that this protocol allows better communication between researchers and facilitates the replication of the models.

[9]Removing the countries that were disconnected from the network.

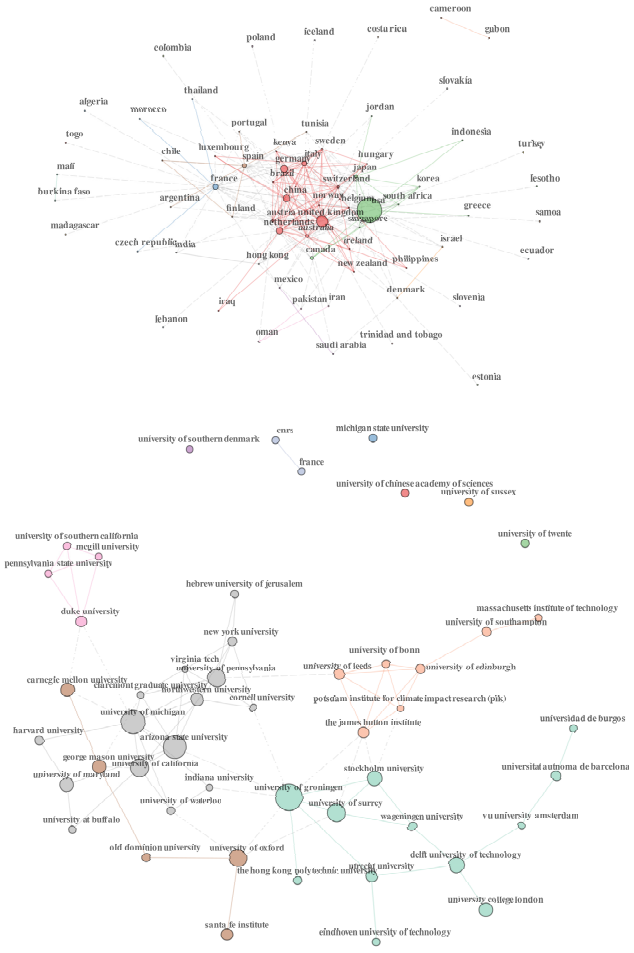

Figure 10: Collaboration networks: between countries (left) and universities (right).

	Countries netw.		Universities netw.	
	ODD	all	ODD	all
Size	50	80	613	1656
Density	0.122	0.077	0.006	0.002
Transitivity	0.476	0.407	0.61	0.447
Diameter	3	4	12	14
Degree Centralization	0.429	0.366	0.054	0.022
Average path length	2.096	2.232	4.372	5.49

Table 6: Network statistics of countries networks and universities networks.

4 Conclusion

The present work carried out a bibliometric analysis of the development of the literature in ABMs in the social sciences and the advancement of the forms of documentation of the ABMs. The area of ABMs is a recently developed area in social sciences in comparison with other important topics, but the study shows an important increase in the last 10 years. This field of research is concentrated both in USA and European countries, but there is a recent development in China and Australia. For the rest of the countries, the research in ABMs is marginal. The importance of documentation for the progress of knowledge about the problems to be analyzed is notorious, although we note that in social sciences it is not central still [10]. The standard for documentation is the ODD protocol, but it has not achieved the degree of penetration that can be seen in other disciplines. Here, we can think that one of the present problems is the tradition of "equation-based" models of the subject. Formalizations such as those involved in ODD (a written-text protocol) are not easy to assimilate in these cases.

The study identifies the most influential journals in the field. In particular, it shows that JASSS is the most influential journal, with the role it plays in the dissemination of articles that use ABMs in social sciences, encouraging articles that follow documentation protocols and its role as a link between articles - and authors - who use ODD compared to others who do not use ODD. In addition, the study provides a detailed analysis of the 150 (and 30) most influential articles, ranked based on their average citations per year. It shows that different topics arise within social simulation, which are not strictly mainstream in any of the disciplines nor are they exclusive to any of them: it is found that the analyzed topics come from multidisciplinary work. These facts are also observed when analyzing keywords co-occurrences and co-citation: heterodox research programs and multidisciplinary work are differential aspects of the analyzed articles.

The degree of use has been uneven according to the different topics to be addressed, with an emphasis on multidisciplinary issues or bordering other disciplines. Most developed areas are the links between ecology and economics, but also analysis of behavior and decision making is a topic of special interest. As we can see, the lines of research that we identify are intrinsically multidisciplinary (feedback between natural and social systems, tourism, energy)[11]. In addition, many of the most important articles are methodological, with an emphasis on reproducibility and simulation analysis. It is clear that here there has been a transfer of knowledge between disciplines, since the ODD protocol, as we have outlined, comes from ecology. However, it is relevant to understand what type of knowledge transfer is happening and how the different disciplines are associated since different configurations imply different approaches to the topics. In this sense, the

[10] We must highlight the efforts of many Journals to improve documentation and transparency in the documentation of the models.

[11] Observe here the pluralistic vision in economics (but that can be extended to the social sciences) where the contribution of other disciplines has allowed continuing the field of study in new directions (see Rodrik (2015); Cedrini and Fontana (2018)).

distinction between interdisciplinary and multidisciplinary approaches to science takes on special relevance. Davis (2016) argues that a relationship between disciplines is necessary where there is a mutual understanding of concepts and meanings, for which it is reasonable to assume that the links will be generated at the boundaries of the disciplines. The link between disciplines must be from this point of view, one of mutual relationship and not one of subordination or appropriation of concepts; the autonomy of disciplines from others may be possible only in the short term. We can therefore think that disciplines and the links between them evolve in a complex manner: the development and current use of ABM documentation protocols are sensitive to how documentation protocols evolved so much in the sciences such as ABMs in social sciences. It is to be expected then that the integration of the social sciences with other disciplines that make intensive use of documentation will allow an improvement in the communication of the models; on the other hand, addressing more issues based on the complexity paradigm will allow, in the long term, the use of these protocols by a larger group of scientists.

Then, we refer to the characterization of the literature that follows the ODD protocol with respect to those that do not. Although we can observe that the research networks involve the same regions, we can affirm that in the topics networks, in the countries networks, and the networks of research centers we can find differences. Networks of researchers working in ODD are denser, which may be an indication of better communication. They also have a higher clustering coefficient and a lower average length path, which can determine a more fluid flow of information in the network.

In conclusion, as the models become more "complex" an adequate description of the multiplicity of factors that affect the model becomes necessary. Is it a necessary condition in multidisciplinary research? We can say from the analysis carried out that there is a greater flow of information in this network and that these networks cover multidisciplinary topics. These differences point to greater multidisciplinary work in the articles that follow the ODD protocol.

We propose some extensions, such as using other databases such as Web of Science, Google Scholar, or preprint repositories, as ARXIV, SSRN, and others. For further analysis, it is also proposed to carry out a comprehensive bibliometric analysis of the modifications that the complex systems approach has caused in the study of the social sciences. In addition, in the future, alternative metrics -including he number of downloads of a paper from specific databases- can be used to measure the impact a study causes, and to provide alternative pictures of the topic.

References

Ahmad, N., Menegaki, A. N., and Al-Muharrami, S. (2020). Systematic literature review of tourism growth nexus: an overview of the literature and a content analysis of 100 most influential papers. *Journal of Economic Surveys*.

An, L. (2012). Modeling human decisions in coupled human and natural systems: Review of agent-based models. *Ecological Modelling*, 229:25–36.

An, L., Zvoleff, A., Liu, J., and Axinn, W. (2014). Agent-based modeling in coupled human and natural systems (chans): lessons from a comparative analysis. *Annals of the Association of American Geographers*, 104(4):723–745.

Arfaoui, N., Brouillat, E., and Saint Jean, M. (2014). Policy design and technological substitution: Investigating the reach regulation in an agent-based model. *Ecological Economics*, 107:347–365.

Aria, M. and Cuccurullo, C. (2017). bibliometrix: An r-tool for comprehensive science mapping analysis. *Journal of informetrics*, 11(4):959–975.

Axelrod, R. (1997). *The complexity of cooperation: Agent-based models of competition and collaboration*, volume 3. Princeton University Press.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the national academy of sciences*, 99(suppl 3):7280–7287.

Bradford, S. C. (1934). Sources of information on specific subjects. *Engineering*, 137:85–86.

Callon, M., Courtial, J.-P., Turner, W. A., and Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Information (International Social Science Council)*, 22(2):191–235.

Cedrini, M. and Fontana, M. (2018). Just another niche in the wall? how specialization is changing the face of mainstream economics. *Cambridge Journal of Economics*, 42(2):427–451.

Davis, J. B. (2016). Economics imperialism versus multidisciplinarity. *History of Economic Ideas*, 24(3):77–94.

Ding, Z., Wang, Y., and Zou, P. X. (2016). An agent based environmental impact assessment of building demolition waste management: Conventional versus green management. *Journal of Cleaner Production*, 133:1136–1153.

Ellegaard, O. and Wallin, J. A. (2015). The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics*, 105(3):1809–1831.

Epstein, J. M. and Axtell, R. (1996). *Growing artificial societies: social science from the bottom up*. Brookings Institution Press.

Farmer, J. D. and Foley, D. (2009). The economy needs agent-based modelling. *Nature*, 460(7256):685–686.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41.

Friege, J., Holtz, G., and Chappin, E. J. (2016). Exploring homeowners' insulation activity. *Journal of Artificial Societies and Social Simulation*, 19(1):4.

Garfield, E. (1988). Announcing the sci compact disk edition-cd-rom gigabyte storage technology, novel software, and bibliographic coupling make desk-top research and discovery a reality. *Current Contents*, (22):3–13.

Garibay, I., Gunaratne, C., Yousefi, N., and Scheinert, S. (2019). The agent-based model canvas: A modeling lingua franca for computational social science. *Social-Behavioral Modeling for Complex Systems*, pages 521–544.

Ge, J. and Polhill, G. (2016). Exploring the combined effect of factors influencing commuting patterns and co2 emissions in aberdeen using an agent-based model. *Journal of Artificial Societies and Social Simulation*, 19(3).

Gilbert, N. (2004). Agent-based social simulation: dealing with complexity. *The Complex Systems Network of Excellence*, 9(25):1–14.

Gilbert, N. (2008). *Agent-based models*, volume 153. Sage Publications.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., et al. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 198(1-2):115–126.

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., and Railsback, S. F. (2010). The ODD protocol: a review and first update. *Ecological Modelling*, 221(23):2760–2768.

Grimm, V., Railsback, S. F., Vincenot, C. E., Berger, U., Gallagher, C., DeAngelis, D. L., Edmonds, B., Ge, J., Giske, J., Groeneveld, J., et al. (2020). The ODD protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism. *Journal of Artificial Societies and Social Simulation*, 23(2).

Hardin, G. (1968). The tragedy of the commons. *Science*, 162:1243–1248.

Hauke, J., Lorscheid, I., and Meyer, M. (2017). Recent development of social simulation as reflected in JASSS between 2008 and 2014: A citation and co-citation analysis. *Journal of Artificial Societies and Social Simulation*, 20(1).

Johnson, P., Nicholls, S., Student, J., Amelung, B., Baggio, R., Balbi, S., Boavida-Portugal, I., de Jong, E., Hofstede, G. J., Lamers, M., et al. (2017). Easing the adoption of agent-based modelling (abm) in tourism research. *Current Issues in Tourism*, 20(8):801–808.

Kessler, M. M. (1962). An experimental study of bibliographic coupling between technical papers. Technical report, MASSACHUSETTS INST OF TECH LEXINGTON LINCOLN LAB.

Lee, J. S., Filatova, T., Ligmann-Zielinska, A., Hassani-Mahmooei, B., Stonedahl, F., Lorscheid, I., Voinov, A., Polhill, G., Sun, Z., and Parker, D. C. (2015). The complexities of agent-based modeling output analysis. *The Journal of Artificial Societies and Social Simulation*, 18(4).

Li, F., Xie, Z., Clarke, K. C., Li, M., Chen, H., Liang, J., and Chen, Z. (2019). An agent-based procedure with an embedded agent learning model for residential land growth simulation: The case study of nanjing, china. *Cities*, 88:155–165.

Macy, M. W. and Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual review of sociology*, 28(1):143–166.

Mantese, G. C. and Amaral, D. C. (2017). Comparison of industrial symbiosis indicators through agent-based modeling. *Journal of Cleaner Production*, 140:1652–1671.

Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., Schlüter, M., Schulze, J., Weise, H., and Schwarz, N. (2013). Describing human decisions in agent-based models–odd+ d, an extension of the odd protocol. *Environmental Modelling & Software*, 48:37–48.

Parker, D. C., Manson, S. M., Janssen, M. A., Hoffmann, M. J., and Deadman, P. (2003). Multi-agent systems for the simulation of land-use and land-cover change: a review. *Annals of the Association of American Geographers*, 93(2):314–337.

Polhill, J. G. (2010). ODD updated. *Journal of Artificial Societies and Social Simulation*, 13(4):9.

Polhill, J. G., Parker, D., Brown, D., and Grimm, V. (2008). Using the ODD protocol for describing three agent-based social simulation models of land-use change. *Journal of Artificial Societies and Social Simulation*, 11(2):3.

Price, D. J. D. S. (1965). Networks of scientific papers. *Science*, pages 510–515.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Richiardi, M. G., Leombruni, R., Saam, N. J., and Sonnessa, M. (2006). A common protocol for agent-based social simulation. *Journal of Artificial Societies and Social Simulation*, 9.

Rodrik, D. (2015). *Economics rules: The rights and wrongs of the dismal science*. WW Norton & Company.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, 1(2):143–186.

Simoes, J. A. (2012). An agent-based/network approach to spatial epidemics. In *Agent-based models of geographical systems*, pages 591–610. Springer.

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for information Science*, 24(4):265–269.

Squazzoni, F. (2010). The impact of agent-based models in the social sciences after 15 years of incursions. *History of Economic Ideas*, pages 197–233.

Squazzoni, F. and Casnici, N. (2013). Is social simulation a social science outstation? a bibliometric analysis of the impact of JASSS. *Journal of Artificial Societies and Social Simulation*, 16(1):10.

Squazzoni, F., Jager, W., and Edmonds, B. (2014). Social simulation in the social sciences: A brief overview. *Social Science Computer Review*, 32(3):279–294.

Vincenot, C. E. (2018). How new concepts become universal scientific approaches: insights from citation network analysis of agent-based complex systems science. *Proceedings of the Royal Society B: Biological Sciences*, 285(1874):20172360.

White, H. D. and McCain, K. W. (1998). Visualizing a discipline: An author co-citation analysis of information science, 1972–1995. *Journal of the American society for information science*, 49(4):327–355.

Wilensky, U. (1999). Netlogo. Evanston, IL: Center for connected learning and computer-based modeling, Northwestern University.

A Appendix A: descriptive statistics of the dataset

Description	Results
MAIN INFORMATION	
Timespan	2007:2020
Sources (Journals, Books, etc)	203
Documents	440
Total Citations	6006
Average years from publication	4.03
Average citations per documents	13.65
Average citations per year per doc	2.147
References	28011
DOCUMENT TYPES	
article	371
book	7
book chapter	28
conference paper	19
editorial	3
letter	1
note	1
review	10
DOCUMENT CONTENTS	
Keywords Plus (ID)	1690
Author's Keywords (DE)	1402
AUTHORS	
Authors	1244
Author Appearances	1524
Authors of single-authored documents	53
Authors of multi-authored documents	1191
Single-authored documents	61
Documents per Author	0.354
Authors per Document	2.83
Co-Authors per Documents	3.46
Collaboration Index	3.14