
Article:
Willett, P. orcid.org/0000-0003-4591-7173 (2022) A bibliometric analysis of the Journal of Molecular Graphics and Modelling: an update. Journal of Molecular Graphics and Modelling, 117. 108313. ISSN 1093-3263

https://doi.org/10.1016/j.jmgm.2022.108313

A Bibliometric Analysis of the *Journal of Molecular Graphics and Modelling*: An Update

Peter Willett[1]

ABSTRACT

This paper provides a bibliometric review of the articles published in the *Journal of Molecular Graphics and Modelling* (formerly the *Journal of Molecular Graphics*). The journal has grown rapidly since its establishment in 1983, with articles coming from countries throughout the world. It now primarily contains articles describing applications of molecular graphics and modelling in chemical and biological systems, rather than the underlying technology that was the focus of many of the early papers in the journal; however, it is these early, system-based papers that continue to attract by far the largest numbers of citations.

KEYWORDS

Bibliometrics, Citation analysis, *Journal of Molecular Graphics and Modelling*, Knowledge exports, Publication statistics

[1] Address: Information School, University of Sheffield, 211 Portobello, Sheffield S1 4DP, UK
Email: p.willett@sheffield.ac.uk

INTRODUCTION

The first issue of the *Journal of Molecular Graphics and Modelling* (or, as it was then called, the *Journal of Molecular Graphics*) appeared in 1983, and 2023 hence sees the 40th anniversary of its foundation. This brief article employs the methods of bibliometrics to illustrate the development of the journal since it was established. Bibliometrics uses publication, citation and - increasingly - download data to investigate, e.g., the impact of the academic research carried out by individuals or institutions, the characteristics of academic networks, or the historical development of a discipline [1-3]. Although first developed in the context of scientific research and based on the *Science Citation Index*, bibliometrics is now being applied across the full range of academe and has several journals specifically devoted to it (e.g., *Scientometrics*, the *Journal of Informetrics* and *Quantitative Science Studies*) in addition to many others that carry frequent bibliometric articles. One application that has found increasing interest is the use of bibliometrics to provide a profile of an individual journal in terms of the articles that it has published and the extent to which those articles have been cited in the literature [4].

The present author and colleagues have published several such profiles of journals in the general area of chemoinformatics and computer chemistry (e.g., [5-8]). The first of these focussed on the *Journal of Molecular Graphics and Modelling* (hereafter JMGM) in a 2007 article that marked 25 years since the founding of its parent body, the Molecular Graphics Society, in 1982. The article studied the changes that had occurred during this quarter-century (as reflected in the articles that had been published in the journal), the most productive and most cited authors and institutions, and JMGM's relationship to other chemoinformatics journals [5]. We now update that report by extending the analysis up to the end of 2021, covering the articles published in the journal that have been indexed in the Clarivate *Web of Science Core Collection* (hereafter WoS) database [9].

ARTICLES IN THE JOURNAL

After the removal of meeting abstracts, editorial material, corrections and the like, a search of WoS in early June 2022 identified a total of 3,237 articles that had been published in JMGM from 1984 to 2021. The set of articles starts with volume 2 (since WoS only started coverage of the journal with that volume) and continues to volume 109, the total of 3,237 including those articles from the 2022 volumes 110 and 111 that had already been published in 2021 as Early Access. The large number of volumes is because the journal moved to single-issue volumes in 2011 with volume 30, whereas the previous volumes had each comprised up to eight individual issues.

The growth in the numbers of published articles is shown in blue in Figure 1, where each point shows the articles appearing in a five-year period (with the exception of the final, 2019-2021 period). The figure also shows the growth in citations, with each orange bar in the bar chart showing the citations

to the journal. It will be seen that the journal has grown consistently over the review period, especially during the last few years: thus, there were 1,684 articles in the first 30 years considered here (1984-2013) and not that many fewer - specifically 1,589 - in just the following eight years (2014-21).

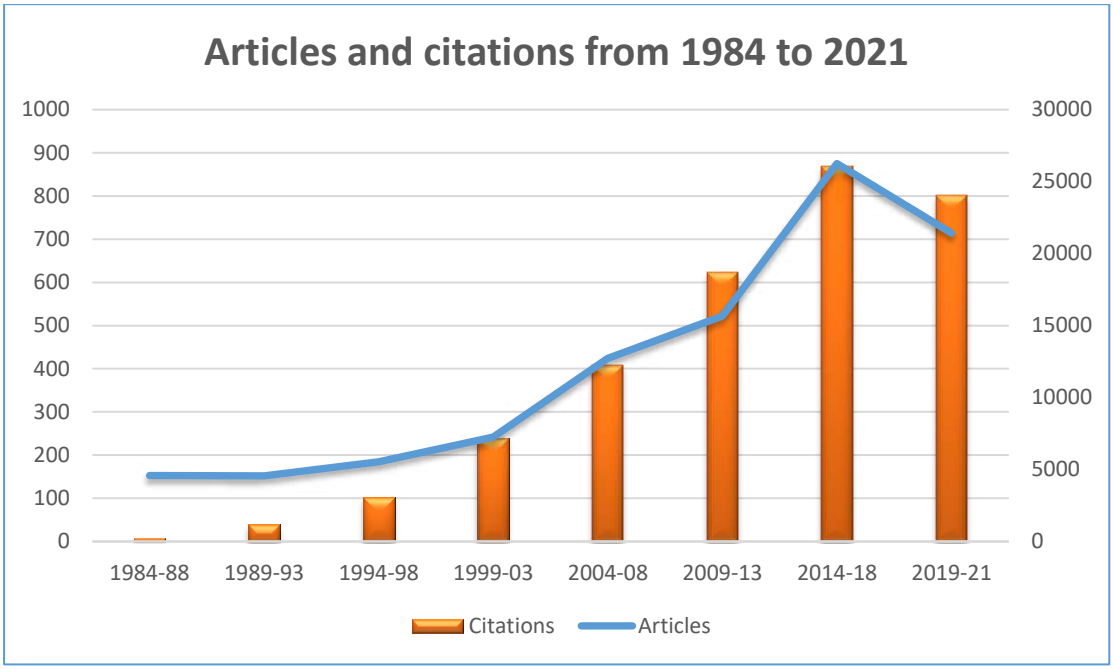


Figure 1. Growth in the articles published in, and in the citations to, the *Journal of Molecular Graphics and Modelling* from 1984 to 2021

The 3,237 articles have authors from 91 different countries. Authors from the USA, China and India have contributed almost one-half of the total number of articles, as shown in the first part of Table 1 which lists the five most productive countries and the numbers of articles that they contributed. However, these overall figures hide the substantial changes that have taken place over time in the source of the articles. Considering the two periods 1984-2013 and 2014-2021, one obtains the rather different sets of counts shown in the second and third parts of the table, where it will be seen that the proportion of USA and UK contributions has declined substantially. This decline is even more striking if one considers just the 538 articles published up to and including 2000, where 322 of them (i.e., 60%) came from these two countries; by way of contrast there were only six contributions from India and only four from the People's Republic of China during this period. The decline reflects the growing internationalization of the subject (as reflected by publication in the journal), since 58 of the 91 nations that have contributed had their first articles published in the current century.

1984-2021		1984-2013		2014-2021	
Nation	Articles	Nation	Articles	Nation	Articles
USA	664	USA	496	China	308
China	484	UK	205	India	293
India	404	China	176	Iran	210
UK	251	India	111	USA	168
Iran	239	Japan	94	Pakistan	75

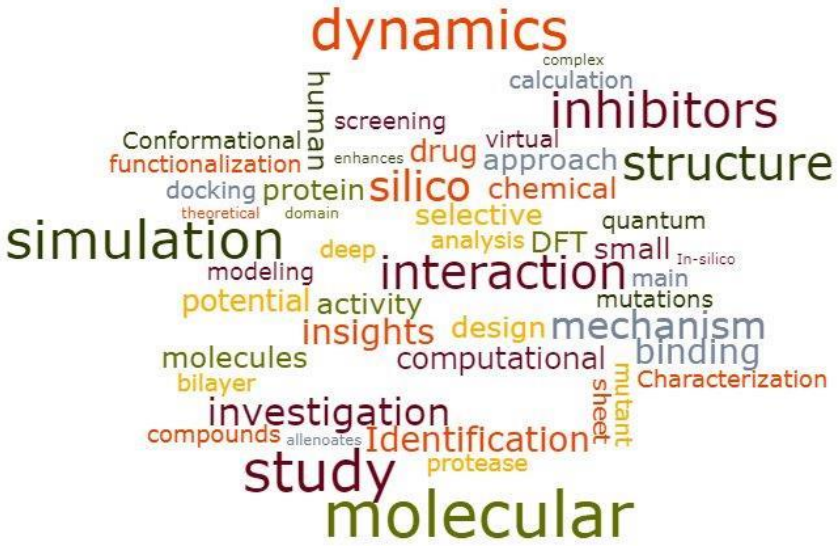
Table 1. Numbers of *Journal of Molecular Graphics and Modelling* articles contributed by the five most productive countries in three time periods

The articles are the work of 8,539 different authors. However, as is normally the case with publication and citation data, the frequency distribution is highly skewed, with no less than 6,668 of these authors contributing only a single article. Conversely, at the other end of the distribution there are seven authors who have contributed 15 or more articles to the journal over the years since it started publication: Khurshid Ayub, Ramanthan Chandiramouli, Serdar Durdagi, Mehdi Esrafili, Noriyuki Kurita, Veerappan Nagarajan, and the present author. Five of these seven authors - Ayub, Chandiramouli, Durdagi, Esrafili, and Nagarajan – were amongst the top seven for the 2014-21 period and yet had contributed just a single article (by Esrafili in 2008) between them prior to 2014, i.e., the journal is seeing a new wave of productive authors as the early pioneers have moved on.

The previous review [5] noted that early issues of the journal were dominated - hardly surprisingly given the journal's title at the time - by descriptions of the details of specific programs or computer systems for graphics and modelling, but that by early in this century the majority of the articles described an increasingly wide range of applications. This has continued to be the case, as is suggested by the words comprising the two word clouds (produced using EdWordle at www.edwordle.net) appearing in Figure 2. The first of these shows the 50 most important words in the titles of the first 100 articles printed in the journal in the period 1984-87, while the second covers the corresponding 50 words from the titles of the last 100 articles published in 2021. While there is a fair degree of overlap between the two, there are fewer words in Figure 2(b) that are directly related to technology, and this is confirmed by inspection of the titles: almost three-quarters of the first-100 articles that are summarised in Figure 2(a) describe novel computational methods, whereas nearly all of the last-100 are applications papers that focus on the analysis of specific chemical or biological entities using some computational technique. Given the rapidly increasing use of machine learning and related techniques in very many branches of science their absence from Figure 2(b) (and also from the journal's scope note on the JMGM website) is perhaps rather surprising. Indeed, a text

search of the 3,237 articles identified only 31 (i.e., less than 1% of them) in response to a search for "machine learning" OR "deep learning", something that will surely change over the next few years.

(a)



(b)

Figure 2. Word clouds of the 50 most important title words from (a) the earliest, and (b) the latest 100 *Journal of Molecular Graphics and Modelling* articles

CITATIONS TO THE JOURNAL

Citations to an article reflect the use that has been made of it by researchers after its publication (or even before publication with the increasing use of preprint systems). There are many different reasons why an article may be cited [10] but it is generally accepted that the number of citations to a specific article provides a measure of its importance to some part of the academic community and of its consequent impact on the development of a subject [11]. The 3,237 JMGM articles had received citations from a total of 84,412 different articles up to the end of 2021, with just 1,546 of them being self-citations (i.e., citations to JMGM by articles that had also been published in JMGM).

Author and title	Reference	Citations
W. Humphey et al., VMD: Visual molecular dynamics	J. Mol. Graph. 14 (1996) 33-38	32,398
R. Koradi et al., MOLMOL: A program for display and analysis of macromolecular structures	J. Mol. Graph. 14 (1996) 51-55	6,179
G.Vriend, WHAT IF - A molecular modeling and drug design program	J. Mol. Graph. 8 (1990) 52-56	3,230
J.M. Wang et al., Automatic atom type and bond type perception in molecular mechanical calculations	J. Mol. Graph. Modell. 25 (2006) 247-260	2,963
A. Golbraikh, and A. Tropsha, Beware of q(2)!	J. Mol. Graph. Modell. 20 (2002) 269-276	2,903
R.M. Esnouf, An extensively modified version of MolScript that includes greatly enhanced coloring capabilities	J. Mol. Graph. Modell. 15 (1997) 132-134	1,807
W. Smith and T.R. Forester, DL_POLY_2.0: A general-purpose parallel molecular dynamics simulation package	J. Mol. Graph. 14 (1996) 136-141	1,705
A.K. Dunker et al., Intrinsically disordered protein	J. Mol. Graph. Modell. 19 (2001) 26-59	1,648
S.V. Evans, SETOR: Hardware-lighted 3-dimensional solid model representations of macromolecules	J. Mol. Graph. 11 (1993) 134-138	1,223
T.E. Ferrin et al., The MIDAS display system	J. Mol. Graph. 6 (1988) 13-27	1,051

Table 2. The most cited articles in the *Journal of Molecular Graphics and Modelling*

Table 2 lists the ten most cited articles, from which it is possible to immediately draw two conclusions: that nearly all of these describe software systems of various sorts, with many of them being the articles that are normally cited when a particular package is used in some piece of research; and that there is an obvious bias towards older articles that have had a greater period of time to attract citations. In these two respects, nothing has changed since the previous JMGM review [5], which made similar comments about the citations to the journal up to 2006, and which are not unexpected given the preponderance of methods articles in early issues of the journal that has been noted above. There is, however, a more striking similarity to the previous review and that is that no less than eight of the articles in Table 2 were also amongst the ten most-cited articles up to 2006, with the only newcomers being the articles by Wang *et al*. and by Golbraikh and Tropsha.

The continuing influence of these early articles is an example of what has become known as the so-called Matthew effect, where an article that has been heavily cited in the past becomes even more cited as time passes [12, 13]. This phenomenon is particularly common with articles describing a novel method: as a new technique - in the case of JMGM the use of a particular software package - starts to be cited in the literature more people become aware of it and then start to use it (often in novel ways as new applications become apparent), this in turn leading to further citations and so on. The growth in citations can be further strengthened if the distributors of a package require authors to cite a particular article when publishing new results that have been obtained from use of the package. These characteristics probably account, at least in part, for the huge number of citations to the article by Humphrey *et al*. at the head of Table 2, its citations alone accounting for 38.3% of all of the citations to JMGM articles, with these citations coming from over 2,000 different journals. The website for the distributors requires "that all published work which utilizes VMD include the primary VMD citation at a minimum", and also lists a large number of other software packages that can be used with VMD [14], something that will again serve to increase citations to it. The distributors of MOLMOL and WHAT IF similarly request users to cite the original JMGM articles when publishing their research. It must be emphasised that highly-cited articles such as these are at the top of an extremely skewed distribution; in contrast to the ten in Table 2, there are 245 articles that have been cited just once and 323 that are uncited (although of course many of these have only been published recently and hence have not yet had the opportunity to attract citations).

The 84,412 citations to JMGM articles came from 4,866 different journals, with 2,221 of these providing just a single citation. Material published in JMGM has hence had an impact, as denoted by citation, for scientists publishing in an extremely wide range of journals. Those contributing the largest numbers of citations are shown in Table 3, which is structured in the same way as Table 1. It will be seen that the *Journal of Physical Chemistry B* and the *Journal of Chemical Information and Modeling* have cited the journal consistently across its lifetime, whereas citations from the *Journal of*

Molecular Biology, *Biochemistry* and the *Journal of Biological Chemistry* have all dropped away very substantially over the last few years. It is interesting to note here the presence of *Scientific Reports*, which first appeared only in 2011. This is an example of an open-access mega-journal (or OAMJ) [15], a relatively new type of publication that has four major characteristics: very large size; completely open-access in operation; a peer-review policy that focusses only on the scientific soundness of a submission's research without considerations of potential impact; and broad disciplinary scope (which for *Scientific Reports* includes all subject areas across the full range of the natural sciences, medicine and engineering). A further OAMJ, *PLOS One* (which was the first OAMJ when it started publication in 2006), provided the ninth largest number of citations in the 2014-2021 period.

1984-2021		1984-2013		2014-2021	
Journal	Citations	Journal	Citations	Journal	Citations
Journal of Physical Chemistry B	2,611	*Journal of Molecular Biology*	1,839	*Journal of Physical Chemistry B*	1,479
Biochemistry	2,141	*Biochemistry*	1,762	*Physical Chemistry Chemical Physics*	1,268
Journal of Molecular Biology	1,995	*Journal of Biological Chemistry*	1,579	*Journal of Biomolecular Structure and Dynamics*	1,084
Journal of Biological Chemistry	1,948	*Journal of Physical Chemistry B*	1,132	*Scientific Reports*	1,005
Journal of Chemical Information and Modeling	1,811	*Journal of Chemical Information and Modeling*	948	*Journal of Chemical Information and Modeling*	863

Table 3. Numbers of citations to the *Journal of Molecular Graphics and Modelling* from the five journals with the largest numbers of citations in three time periods. The citations from *Journal of Chemical Information and Modeling* include those from its predecessor, the *Journal of Chemical Information and Computer Sciences*

The presence of OAMJs highlights a limitation of Table 3, in that it takes no account of the sizes of the journals, since one might expect the number of citations to JMGM to be related, in part at least, to the number of articles in the citing journal. Thus the *Journal of Physical Chemistry B*'s 2,611 citations came from 43,097 articles that were published there from 1984 to 2021, while the *Journal of Chemical Information and Modeling*'s 1,811 citations came from just 7,797 articles during the same period. The latter journal would hence appear to be much the more strongly related to JMGM, as might be expected given that many of its articles cover similar ground to that published in JMGM. The previous review [5] discussed the citations between JMGM and several other journals in the general areas of chemoinformatics and molecular modelling, using relatedness data from the *Journal Citation Reports* database. This data is no longer available, but a simple measure of the strength of the relationship between JMGM and another journal, *A*, for the period 1984-2021 can be obtained by normalizing the number of citations from *A* to JMGM by the number of articles in *A* for that period. This gives the following values for five cognate journals: *Journal of Computer-Aided Design* 0.29; *Journal of Chemical Information and Modeling* (including its predecessor, the *Journal of Chemical Information and Computer Sciences*) 0.23; *Molecular Informatics* (including its predecessors *Quantitative Structure-Activity Relationships* and *QSAR and Combinatorial Science*) 0.20; *Journal of Molecular Modeling* 0.19; and *Journal of Cheminformatics* 0.16. The *Journal of Computer-Aided Molecular Design* would hence appear to have the strongest citation linkage to JMGM (as it did based on relatedness data in the previous review); at 0.46, the value for JMGM itself is, hardly surprisingly, larger than any of the values above.

KNOWLEDGE EXPORTS

A more detailed analysis of the journal's impact, as denoted by the citations to its articles, can be obtained from consideration of the WoS subject categories to which the citing articles belong, since every article is assigned to one or more of 254 different categories when it is entered into the database.

The great majority of the citations to JMGM come from journals such as those discussed thus far; however, others come from journals in subject categories that are very different from the many categories associated with chemistry, biology and computation that one would expect to see. Thus, the ten largest contributions, providing a total of 67,917 of the 84,412 citations, come from the following cognate categories (in order of decreasing numbers of citations): Biochemistry Molecular Biology; Chemistry Physical; Chemistry Multidisciplinary; Biophysics; Physics Atomic Molecular Chemical; Chemistry Medicinal; Computer Science Interdisciplinary Applications; Materials Science Multidisciplinary; Multidisciplinary Sciences; and Biochemical Research Methods. However, citations come from a further 192 categories, many of which would seem, on first appearance at least, to have little or nothing to do with JMGM's subject matter. Citations such as these provide rather

unexpected examples of what can be regarded as knowledge exports, i.e., the transfer of knowledge from one academic field (as denoted here by a WoS category) to a different field [16, 17]. Examples here include citations to articles in Table 2 that come from articles covering topics as diverse as: risk factors for the West Nile virus (an article that had been assigned to the Public, Environmental & Occupational Health, and to the Social Sciences Biomedical categories) [18]; the efficiency of the Algerian banking system (Management) [19]; the origins of life within the theoretical frameworks of thermodynamics, informatics and self-reproducing automata (Humanities Multidisciplinary, and History & Philosophy of Science) [20]; and the design of a system for experiencing virtual environments and performances (Behavioral Sciences, Computer Science Information Systems, Computer Science Theory & Methods, Music, Neurosciences, and Psychology Experimental) [21]. It must be emphasised that this behaviour is not specific to JMGM in any way since other such examples are described in analyses of citations to articles in the *Journal of Mathematical Chemistry* [6] and to articles about the Cambridge Crystallographic Database [7]; instead, it simply demonstrates the widespread impact of the research published in the journal.

CONCLUSIONS

This brief paper has summarized the publications in, and the citations to articles in, the *Journal of Molecular Graphics and Modelling*. There is a fair degree of continuity with the conclusions of the previous bibliometric review in 2007 [5], which noted that the journal's most highly cited papers were those describing systems or algorithms, but that the proportion of these papers was decreasing as more applications of molecular graphics and molecular modelling started to be developed. This is clearly still the case, as described above; however, the final conclusion in 2007 was that while the journal was international in scope, the USA provided by far the largest number of articles. It remains international in scope, indeed with noticeably more nations contributing now than in 2007, but the previous domination by the USA (and, to a lesser extent, the UK) is a thing of the past as a new generation of researchers from around the world have started to make their contributions to this important area of science.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CONFLICTS OF INTEREST

The author declares no competing interests.

REFERENCES

1. J. Mingers and L. Leydesdorff, A review of theory and practice in scientometrics. Eur. J. Operat. Res. 246 (2015) 1-19. DOI: 10.1016/j.ejor.2015.04.002
2. P. Wouters, M. Thelwall, K. Kousha et al., The Metric Tide: Literature Review. DOI: 10.13140/RG.2.1.5066.3520. At https://www.ukri.org/publications/review-of-metrics-in-research-assessment-and-management/
3. D.W. Aksnes, L. Langfeldt, and P. Wouters, Citations, citation indicators, and research quality: An overview of basic concepts and theories. SAGE Open 9:1 (2019). DOI: 10.1177/2158244019829575.
4. K.W.U. Anyi, A.N. Zainab, and N.B. Anuar, Bibliometric studies on single journals: a review. Malay. J. Libr. Inform. Sci. 14 (2009) 17–55.
5. P. Willett, A bibliometric analysis of the *Journal of Molecular Graphics and Modelling*. J. Mol. Graphics Modell. 26 (2007) 602–606. DOI: 10.1016/j.jmgm.2007.03.008
6. G. Restrepo and P. Willett, The *Journal of Mathematical Chemistry*: a bibliometric profile. J. Math. Chem. 55 (2017) 1589–1596. DOI: 10.1007/s10910-017-0747-7
7. P. Willett, J.C. Cole, and I.J. Bruno, The impact of the Cambridge Structural Database and the small molecule crystal structures it contains: a bibliographic and literature study. CrystEngComm 22 (2020) 7233–7241. DOI: 10.1039/d0ce00045k
8. P. Willett, The literature of chemoinformatics, 1978-2018. Int. J. Mol. Sci. 21 (2020) 5576. DOI: 10.3390/ijms21155576
9. C. Birkle, D.A. Pendlebury, J. Schnell et al., Web of Science as a data source for research on scientific and scholarly activity. Quant. Sci. Stud. 1 (2020) 363–376. DOI: 10.1162/qss_a_00018
10. L. Bornmann and H. Daniel, What do citation counts measure? A review of studies on citing behaviour. J. Docum. 64 (2006) 45-80. DOI: 10.1108/00220410810844150
11. L. Waltman (2016). A review of the literature on citation impact indicators. J. Informet. 10 (2016) 365-391. DOI: 10.1016/j.joi.2016.02.007
12. R.K. Merton, The Matthew effect in science: the reward and communication systems of science are considered. Science 159(3810) (1968) 56–63.
13. M. Perc, The Matthew effect in empirical data. J. Roy. Soc. Interface11 (2014) 20140378. DOI: 10.1098/rsif.2014.0378
14. NIH Center for Macromolecular Modeling and Bioinformatics, University of Illinois at Urbana-Champaign, VMD - Visual Molecular Dynamics. At https://www.ks.uiuc.edu/Research/vmd
15. V. Spezi, S. Wakeling, S. Pinfield et al., Open-access mega-journals: The future of scholarly communication or academic dumping ground? A review. J. Docum. 73 (2017) 263-283. DOI: 10.1108/JD-06-2016-0082

16. A. Lockett and A. McWilliams, The balance of trade between disciplines: Do we effectively manage knowledge? J. Manag. Inquiry 14 (2005) 139-150. DOI: 10.1177/1056492605276645
17. E. Yan, Y. Ding, B. Cronin B et al., A bird's-eye view of scientific trading: dependency relations among fields of science. J Informetr. 7 (2013) 249–264. DOI: 10.1016/j.joi.2012.11.008
18. D. Ghosh and R. Guha, Using a neural network for mining interpretable relationships of West Nile risk factors. Soc. Sci. Med.72 (2011) 418-429. DOI: 10.1016/j.socscimed.2010.09.014
19. A. Achi, Efficiency and its determinants in the Algerian banks: network data envelopment analysis and partial least squares regression. Int. J. Product. Perform. Manag. in press. DOI: 10.1108/IJPPM-07-2021-0407
20. P.R. Wills, Genetic information, physical interpreters and thermodynamics; the material-informatic basis of biosemiosis. Biosemiot. 7 (2014) 141-165. DOI: 10.1007/s12304-013-9196-2
21. X. Amatriain, J. Castellanos, T. Höllerer. Experiencing audio and music in a fully immersive environment. Lect. Notes Comput. Sci. 4969 (2008) 380–400. DOI: 10.1007/978-3-540-85035-9_27