Field classification in Dimensions: A case study of textile technology

Jokić, Davor

Field classification in Dimensions: A case study of textile technology

Davor JOKIĆ
Faculty of Textile Technology, Library, University of Zagreb, Zagreb, Croatia
davor.jokic@ttf.hr

Original scientific article
UDC 001:677
DOI: 10.31881/TLR.2019.31
Received 20 May 2019; Accepted 8 July 2019; Published Online 10 July 2019

ABSTRACT

Given the latest research on Dimensions classification, this article discusses the novelty of such classification in the field of textile technology from the standpoint of Croatian scientific career advancement system. New machine learning article based classification system is compared to a traditional journal based classification system brought by the Web of Science and Scopus in terms of evaluation significance. The starting point of assigned category comparison were 13 journals indexed in the Web of Science in just one common category - Materials science, Textiles. Since Scopus does not have a unique category for the textile technology a list of 11 assigned categories was put in the comparison. Lastly, 58 research fields assigned to the articles published in mentioned journals indexed in Dimensions were analyzed for validity. Results show that the unique category of Textiles in Web of Science fully fits the field of textile technology from Croatian point of view. Scopus model with multi category assignment is not so reliable and useful in field evaluation. Lastly, Dimensions with its novel approach failed to validly classify indexed publications.

KEYWORDS

Dimensions, bibliometrics, field classification, textile technology, research evaluation

INTRODUCTION

The recent studies on relatively new research data platform Dimensions produced sufficient amount of information about its origin and features [1,2,3]. The main characteristics can be pointed out for the first time users. In 2018 a technology company Digital Science launched their product named Dimensions[1] offering scientific and research communities a research data infrastructure and tool without charge. The purpose was to introduce a new way the research is discovered, accessed and analyzed.

The main novelty is a classification system based on machine learning automatic category assignment on the level of single publication. Such classification is completely different from traditional journal based classification brought by relevant scientific databases, especially the Web of Science and Scopus. The reason for employment of such model was an effort to solve a problem of consistent categorization of grants, patents and clinical trials, indexed by Dimensions, which could not be identified within journal-based category system [4].

[1] www.dimensions.ai

In addition, Bornmann implies an advantage of recognition of multidisciplinary journals, which are not classified in corresponding fields in the context of standard journal-based scheme [3]. It is supposed that single paper classification should give a more precise picture of multidisciplinary journals. Instead of one to few categories assigned to a journal, Dimensions offers a large amount of categories, of which few, assigned to single paper. In that way the categories with the biggest share in total number of categories assigned to the papers published in the indexed journal should point out the journal scope.

In that regard, this paper brings further research on the novelty of Dimensions classification system in terms of applicability in evaluation of the field of textile technology from the standpoint of Croatian scientific career advancement system. The textile technology was used as a case study because it is the author's field of interest and professional activity.

The Croatian scientific career advancement system and overall evaluation of science is oriented on journal-based classification system. Science is classified in to the fields and subfields according to the Fields of Science and Technology (FOS) published by Organization for Economic Co-operation and Development (OECD). In order to advance in higher title, a researcher or academic must meet a number of criteria among which are papers published in a journal within the same field of science as his or hers vocation. The same is with evaluation and accreditation of academic or scientific institutions. Members of such institutions are supposed to publish papers in the journals with the aim and scope of the same scientific field as its institution's registered activity.

Given that, the usefulness of the introduced novelties of Dimensions classification will depend on accuracy of machine learning category assignment opposed to the traditional journal based classification system used by the Web of Science and Scopus.

Assuming that the Croatian model is not the only one and that there are some similar models in other countries an analysis of possible benefits of Dimensions classification in such a context is conducted in order to get a bigger picture.

Initial studies already put in question the reliability and general validity of results of machine learning classification at the article level. Although, it is likely to be questionable in the field of Textile technology as well, this research will demonstrate empirically the accuracy of this classification. Since Dimensions aims to offer the alternative to traditional classification, it is important to conduct such a research. Users from textile research community that are not familiar with the new Dimensions platform should have an insight of its value. In addition, previous articles on this matter stated that more empirical studies are needed [1,2,3].

Within the novelty of Dimensions platform, another question arises. In the paper bibliographic records one key element is missing – the keywords. Could it be assumed that the assigned categories took over the role of the keywords in some way? If a comparison is made, the purpose of the keywords is virtually the same as the assigned research categories. The point is to describe the specific paper with the standardized words enabling the interconnection in terms of discovering related papers or determination of specific field of science. The absence of the keywords in the Dimensions paper bibliographic records, while both Web of Science and Scopus employ them, goes in favor of the assumed. Besides the author keywords, stated in the papers themselves, Web of Science and Scopus offer service like KeyWords Plus and Indexed keywords bringing the additional relevant keywords that were not listed by the author or publisher. These services offer the greater possibility of uncovering more papers that may not have been listed in the results of one's search [5].

The real reason behind the absence of the keywords is unknown. Additionally surprising is the fact that the keywords are stated in the papers anyhow so it would be easy to index them and present in the paper record.

The reasons for the lack of the keywords should be investigated further or should be explained by Dimensions. If the argument of relationship between the field categories and keywords is valid, the importance of introduced novelties is reduced.

METHOD

The aim of this research is to see whether the novelty of Dimensions classification has some usability value in comparison to standard journal based classification systems brought by relevant scientific databases in terms of journals and papers they index within the field of textile technology. If so, assigned categories should match with those assigned by Web of Science and Scopus. Additionally, in what proportion categories assigned by Dimensions relate to the journals identified to pertain to the field of textile technology.
Methodological approach time span of the research was 2017 since it is the last year with available data in the InCites Journal Citation Reports (JCR) and Scimago Journal & Country Rank (SJR). There was no need to use a wider time span than a year, because it gave representative number of papers (N=1163) and if there were some errors or misclassifications in the data they would appear most likely. Comparison of the results in the Web of Science, Scopus and Dimensions is made, analyzing the results for any correlation.
The starting point of the research was the Web of Science category Materials Science, Textiles in the JCR. This category was chosen since it gathers all indexed journals focused on the manufacture of clothing and furniture from materials made of natural fibers (e.g., leather, cotton, wool, wood) and/or synthetic fibers (e.g., polyester, vinyl, nylon), covering dyes and colors and fiber chemistry [6]. There are 24 indexed journals in this category in 2017. The 13 of them have the Materials Science, Textiles as only assigned category. For the reason of strict determination of textile technology, these 13 journals will be analyzed and compared in Scopus and Dimensions classification. Other 11 journals have few categories assigned one of which is Materials Science, Textiles. They are excluded for the possibility of misleading in direction of other assigned categories. Shown in the Table 1.

Table 1. Journals indexed in Web of Science category Materials Science, Textiles in InCites Journal Citation Reports in 2017

InCites Journal Citation Reports — Clarivate Analytics

Journal Data Filtered By: Selected JCR Year: 2017 Selected Editions: SCIE Selected Categories: 'MATERIALS SCIENCE, TEXTILES' Selected Category Scheme: WoS

Rank	Full Journal Title
1	CELLULOSE
2	DYES AND PIGMENTS
3	**TEXTILE RESEARCH JOURNAL** (Only - Materials Science, Textiles)
4	FIBERS AND POLYMERS
5	**Journal of Industrial Textiles** (Only - Materials Science, Textiles)
6	**JOURNAL OF THE TEXTILE INSTITUTE** (Only - Materials Science, Textiles)
7	COLORATION TECHNOLOGY
8	JOURNAL OF VINYL & ADDITIVE TECHNOLOGY
9	**Journal of Natural Fibers** (Only - Materials Science, Textiles)
10	**Autex Research Journal** (Only - Materials Science, Textiles)
11	WOOD AND FIBER SCIENCE

12	JOURNAL OF THE AMERICAN LEATHER CHEMISTS ASSOCIATION
13	**Journal of Engineered Fibers and Fabrics** (Only - Materials Science, Textiles)
14	Journal of Fiber Science and Technology
15	**FIBRES & TEXTILES IN EASTERN EUROPE** (Only - Materials Science, Textiles)
16	**International Journal of Clothing Science and Technology** (Only - Materials Science, Textiles)
17	**JOURNAL OF THE SOCIETY OF LEATHER TECHNOLOGISTS AND CHEMISTS** (Only - Materials Science, Textiles)
18	**Industria Textila** (Only - Materials Science, Textiles)
19	**INDIAN JOURNAL OF FIBRE & TEXTILE RESEARCH** (Only - Materials Science, Textiles)
20	**AATCC Journal of Research** (Only - Materials Science, Textiles)
21	AATCC REVIEW
22	FIBRE CHEMISTRY
23	**Tekstil ve Konfeksiyon** Only - Materials Science, Textiles)
24	SEN-I GAKKAISHI

*Web of Science InCites Journal Citation Reports, institutional access by AAI@EduHr – Croatian Research and Education Federation

The second stage of the research was to analyze Scopus based SJR categories assigned to those 13 journals for any correlation. Data for Journal of Engineered Fibers and Fabrics are not available so 12 journal will be compared for relations in assigned categories.
Unfortunately, SJR classification does not have a specific category related to the textile technology, so the results show 11 categories assigned to the journals. Shown in Chart 1.

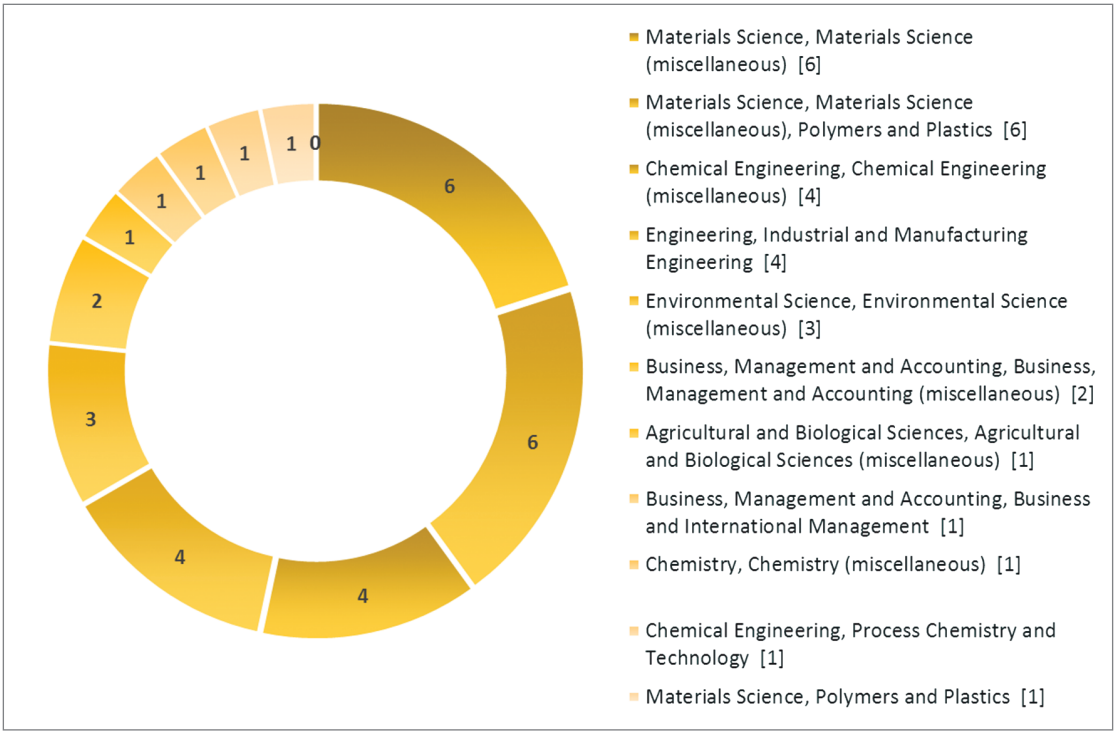

Chart 1. Representation of SJR subject areas and categories per number of journals in research

Besides the lack of textile technology category, it can be seen, from the Chart 1., that there is not one unique category that is assigned to all 13 journals. It can also be seen that there are some overlapping categories. If related categories are put together, we get that every journal is assigned with category of Materials Sciences. This subject area is too broad to define anything specific. Also, it is the only connection between

these journals, besides possible citation relations whose analysis is not conducted in this research, showing the limited possibilities of field research in Scopus.

RESULTS

It can be said that small group of journals (N=13) indexed in Scopus with scope of textile technology has an unnecessary dispersion in 7 subject areas or 11 subject categories making it impossible to compare them on category level. In addition, there is one deviation. The subject area of Agricultural and Biological Sciences is assigned to The Journal of The Textile Institute. Stated in the journal's scope it welcomes papers concerning research and innovation, reflecting the professional interests of the Textile Institute in science, engineering, economics, management and design related to the textile industry and the use of fibers in consumer and engineering applications [7]. The reason for misclassification is probably, as Wang & Waltman state, in Scopus inaccurate classification [8].

The 10 SJR categories are valid in describing major aspects of textile technology but only when reached through targeted journals, which requires advanced knowledge. In fact, these general categories are useless for not being able to present the group of journals that are related to each other. In this sense, the absence of the specific category of the textile technology is obvious fault.

Lastly, the analysis of categories assigned to the single papers of all 13 journals in Dimensions is not possible because four journals do not have their papers classified. The reason is unknown, but it can be assumed that it is the consequence of Dimensions product development and that those papers will be classified eventually. The nine journals with categories assigned to the single papers show the total of 58 different research fields (Table 2. and Table 3.).

Table 2. Journals with single paper assigned categories indexed in Dimensions out of 13 journals in the comparison

No.	Journal title
1	AATCC JOURNAL OF RESEARCH
2	AUTEX RESEARCH JOURNAL
3	FIBRES & TEXTILES IN EASTERN EUROPE
4	INTERNATIONAL JOURNAL OF CLOTHING SCIENCE AND TECHNOLOGY
5	JOURNAL OF ENGINEERED FIBERS AND FABRICS
6	JOURNAL OF INDUSTRIAL TEXTILES
7	JOURNAL OF NATURAL FIBERS
8	JOURNAL OF THE TEXTILE INSTITUTE
9	TEXTILE RESEARCH JOURNAL

Dimensions implements the Field of Research (FOR) system covering all areas of research from the Australian and New Zealand Standard Research Classification (ANZSRC) using automated allocation of FOR codes to indexed documents based on category definitions defined by machine learning [9]. The ANZSRC has a field of textile technology in its classification. It is under division 09 Engineering; group 0910 Manufacturing Engineering as 091012 Textile Technology. Along that, there is a textile engineering under the 0912 Materials Engineering within the same division 09 Engineering. However, analysis of categories assigned to the nine journals shows that there is not a single paper with Textile Technology assigned. The reason for that lies in the Dimensions categorization that is on level 2 with 4 digit codes unlike the original Field of Research

system, which has three levels with 2, 4 and 6 digit codes. The field 091012 Textile Technology is on the third level with six-digit code.

As such, the complete three level Australian and New Zealand Standard Research Classification (ANZSRC) has 22 divisions, 157 groups and 1238 fields. For comparison, the Web of Science has 235 categories and Scopus uses 27 subject areas with 287 categories. Given that, it can be assumed that Dimensions nonuse of the third level of FOR classification has something to do with the large number of fields. The real reason is unknown.

Table 3. Dimensions categories in relation to number of journals to which papers they are assigned to

Category	Number of journals
Manufacturing Engineering 0910	9
Materials Engineering 0912	9
Physical Chemistry (incl. Structural) 0306	9
Artificial Intelligence and Image Processing 0801	7
Civil Engineering 0905	7
Statistics 0104	7
Clinical Sciences 1103	6
Environmental Engineering 0907	6
Interdisciplinary Engineering 0915	6
Applied Mathematics 0102	5
Biochemistry and Cell Biology 0601	5
Macromolecular and Materials Chemistry 0303	5
Other Physical Sciences 0299	5
Psychology 1701	5
Applied Economics 1402	4
Biomedical Engineering 0903	4
Business and Management 1503	4
Chemical Engineering 0904	4
Information Systems 0806	4
Microbiology 0605	4
Analytical Chemistry 0301	3
Cardiorespiratory Medicine and Haematology 1102	3
Communications Technologies 1005	3
Electrical and Electronic Engineering 0906	3
Inorganic Chemistry 0302	3
Medical Physiology 1116	3
Numerical and Computational Mathematics 0103	3
Plant Biology 0607	3
Atomic, Molecular, Nuclear, Particle and Plasma Physics 0202	2
Genetics 0604	2
Nanotechnology 1007	2
Public Health and Health Services 1117	2
Pure Mathematics 0101	2
Soil Sciences 0503	2

Archaeology 2101	1
Cognitive Sciences 1702	1
Computer Software 0803	1
Crop and Pasture Production 0703	1
Demography 1603	1
Design Practice and Management 1203	1
Econometrics 1403	1
Geophysics 0404	1
Historical Studies 2103	1
Human Movement and Sports Science 1106	1
Language Studies 2003	1
Law 1801	1
Linguistics 2004	1
Marketing 1505	1
Optical Physics 0205	1
Organic Chemistry 0305	1
Other Chemical Sciences 0399	1
Other Medical and Health Sciences 1199	1
Performing Arts and Creative Writing 1904	1
Physiology 0606	1
Resources Engineering and Extractive Metallurgy 0914	1
Sociology 1608	1
Specialist Studies In Education 1303	1
Veterinary Sciences 0707	1

At first glance, the distribution of the valid categories on journal level is 100%. All 10 journals have Manufacturing Engineering and Materials Engineering assigned to at least one paper. Both group categories include textile technology and textile engineering outlining the field of investigation. Third and last category that is assigned to at least one paper published in every journal is Physical Chemistry, which also have meaningful relation.

Nevertheless, the results on the single paper level are not so favorable. In total number of papers (N=918) published in nine journals in 2017, categories Manufacturing Engineering and Materials Engineering are underrepresented with 143 (15.6%) and 298 (32.4%) papers. Such a small share of valid categories is completely unexpected. It is not even possible to define the related field of science to the journals. With such high dispersion of categories and obvious incorrectly assigned categories whole classification model comes in question. The worst thing is that group 0910 Manufacturing Engineering, to which 091012 Textile Technology strictly belongs in FOR system, has second smallest share of three most represented categories. The share of 15.6% is significantly smaller than Materials Engineering (32.5%), which does not even have field related to the textile technology in its classification. Textile engineering is only mentioned in the context of 091209 Polymers and Plastics. Representations are shown in Table 4.

Table 4. Representation of relevant categories per indexed papers

Category	Number of papers*	Share
Materials Engineering 0912	298	32.5%
Manufacturing Engineering 0910	143	15.6%
Physical Chemistry (incl. Structural) 0306	137	14.9%

*Total number of papers - 918

In addition to that, in total relations of all indexed papers, regardless of years published, in nine journals the picture is even worse. Manufacturing Engineering has a share of 12.5%, Materials Engineering has 17.2% and Physical Chemistry has 11.6%.

The most of the other categories point out various aspects of textile technology but only under the condition that those categories are not the assigned primary. They can be assigned as a complementation to the field of Manufacturing Engineering.

Further analysis of currently assigned categories shows obvious mistakes. For example, article entitled *Characterization of Thermal Properties of Pig Hair Fiber* published in Journal of Natural Fibers (2017, Vol. 14, Iss. 2) is classified as 0707 Veterinary Sciences. The field of research is incorrectly assigned because article investigates the thermal properties of pig hair fiber in order to provide insights for application in places where natural fibers are utilized for insulation. Another article entitled *Using a 3D Body Scanner in Designing Compression Products Supporting External Treatment* published in Fibres and Textiles in Eastern Europe (2017, Vol. 25, Iss. 5) shows the same inaccuracy. The article comprises a statistical tolerance analysis of human body dimensions using a 3D body scanner and its impact on the value of unit pressure exerted by a compression product on the subject's body. Instead of clothing technology or its related category the assigned field of research is 1801 Law. The reason for that is the Laplace law used in model calculations of changes in unit pressure. The machine picked the term *law* and automatically assigned it to the category of 1801 Law. The same principle repeats throughout the majority of indexed papers. The list of articles with incorrectly assigned categories is too long to present them all. These two examples are quite enough to get the insight. It is impossible to conduct deeper analysis because there are no valid grounds. Application of advanced bibliometric methods would be pointless.

The FOR system has classified Textile Technology as a field of research so a suggestion goes to the Dimensions to employ it. If the inherent lack of the FOR's third level of classification in Dimensions is put aside, still the under-representation of Manufacturing Engineering indicates inaccuracy of classification model. Thereby, the Dimensions cannot be recommended for valid analysis, interpretations or result based decision making in terms of field of textile technology. In that regard the possibility of free use, compared to Web of Science and Scopus subscription, is insignificant.

CONCLUSION

Comparison of traditional journal based classification, brought by relevant scientific databases, and classification system based on machine learning automatic category assignment on the level of single publication showed some significant differences.

The first level of marked differences is the single paper classification, which disables the journal search by scientific fields. This kind of search is typical for users who wish to explore their possibilities of publication in specific field of research. The comparison of journals, by quartiles according to metric indicators like

impact factor or SJR, is also common. Due to the nonuse of the third level of FOR's classification Dimensions provides too broad and general fields for any kind of specific research. Standard metric indicators are also unavailable, at least in the free version of Dimensions which was used in this research.

The second level of mentioned differences is related to the machine learning automatic category assignment. It just does not classify papers correctly. There are numerous misclassifications, which could lead to the wrong conclusions. All papers, which are not assigned with the Manufacturing Engineering, are in the context of this research incorrect. Currently assigned categories can only be complementary or additional in terms of outlining the aspect of research being conducted. The core remains in the field of textile technology. Results show that the unique category of Textiles in Web of Science fully fits the field of textile technology from Croatian point of view. Dimensions machine learning automatic category assignment on the level of single publication has a potential for providing valuable information but in reality, it just does not work right. Application of the third level of FOR's classification and expert guided machine learning is suggested. That is the only way that Dimensions could offer the alternative and rival the Scopus and Web of Science in some way.

REFERENCES

[1] Thelwall M. Dimensions: A competitor to Scopus and the Web of Science?. Journal of Informetrics [Internet]. 2018 Mar [cited 2019 Apr 25];12(2):430-435. Available from: https://www.sciencedirect.com/science/article/abs/pii/S175115771830066X doi: 10.1016/j.joi.2018.03.006

[2] Orduna-Malea E, Delgado Lopez-Cozar E. Dimensions: re-discovering the ecosystem of scientific information. El Profesional de la Informacion [Internet]. 2018 Apr [cited 2019 Apr 25];27(2):420-431. Available from: https://arxiv.org/abs/1804.05365 doi: 10.3145/epi.2018.mar.21

[3] Bornmann L. Field classification of publications in Dimensions: a first case study testing its reliability and validity. Scientometrics [Internet]. 2018 Jul [cited 2019 Apr 25];117(1):637-640. Available from: https://link.springer.com/article/10.1007/s11192-018-2855-y doi: 10.1007/s11192-018-2855-y

[4] Bode C, Herzog C, Hook D, McGrath R. A Guide to the Dimensions Data Approach - A collaborative approach to creating a modern infrastructure for data describing research: where we are and where we want to take it [Internet]. Cambridge: Digital Science; 2018 [cited 2019 Apr 25]. 24 p. Available from: https://www.digital-science.com/resources/portfolio-reports/a-guide-to-the-dimensions-data-approach/

[5] THOMSON REUTERS. Web of Knowledge: User tips – research made easy [Internet]. 2010 [Cited 2019 Apr 26]. Available from: http://interest.science.thomsonreuters.com/content/WOKUserTips-201010-IN

[6] Clarivate Analytics. InCites Journal Citation Reports: Materials Science, Textiles – Category Profile [Internet]. 2019 [Cited 2019 May 8]. Available from: https://jcr.clarivate.com/JCRCategoryProfileAction.action?year=2017&categoryName=MATERIALS%20SCIENCE%2C%20TEXTILES&edition=SCIE&category=QJ

[7] Taylor & Francis Online. The Journal of The Textile Institute: Aims and scope [Internet]. 2019 [Cited 2019 May 10]. Available from: https://www.tandfonline.com/action/journalInformation?show=aimsScope&journalCode=tjti20

[8] Wang Q, Waltman L. Large-scale analysis of the accuracy of the journal classification systems of Web of Science and Scopus. Journal of Informetrics [Internet]. 2016 [Cited 2019 May 10]. Available from: https://www.sciencedirect.com/science/article/abs/pii/S1751157715301930 doi: 10.1016/j.joi.2016.02.003

[9] Digital Science. Dimensions – Fields of Research [Internet]. 2019 [Cited 2019 May 13]. Available from: https://app.dimensions.ai/browse/publication/for?and_facet_year=2017&and_facet_source_title=jour.1136310&redirect_path=/analytics/publication/for/aggregated