Controlled Vocabularies and Information Retrieval: 1918 Pandemic's Scientific Literature as an Example

M. Garcia-Alsina, J. Cobarsí

Abstract—The role of controlled vocabularies in information retrieval is broadly recognized as a relevant feature. Besides, there is a standing demand that editors and databases should consider the effective introduction of controlled vocabularies in their procedures to index scientific literature. That is especially important because information retrieval is pointed out as a significant point to drive systematic literature review. Hence, a first question emerges: Are the controlled vocabularies at this moment considered? On the other hand, subject searching in the catalogs is complex mainly due to the dichotomy between keywords from authors versus keywords based on controlled vocabularies. Finally, there is some demand to unify the terminology related to health to make easier the medical history exploitation and research. Considering these features, this paper focuses on controlled vocabularies related to the health field and their role for storing, classifying, and retrieving relevant literature. The objective is knowing which role plays the controlled vocabularies related to the health field to index and retrieve research literature in data bases such as Web of Science (WoS) and Scopus. So, this exploratory research is grounded over two research questions: 1) Which are the terms considered in specific controlled vocabularies of the health field; and 2) How papers are indexed in relevant databases to be easily retrieved, considering keywords vs specific health' controlled vocabularies? This research takes as fieldwork the controlled vocabularies related to health and the scientific interest for 1918 flu pandemic, also known equivocally as 'Spanish flu'. This interest has been fostered by the emergence in the early 21st of epidemics of pneumonic diseases caused by virus. Searches about and with controlled vocabularies on WoS and Scopus databases are conducted. First results of this work in progress are surprising. There are different controlled vocabularies for the health field, into which the terms collected and preferred related to '1918 pandemic' are identified. To summarize, 'Spanish influenza epidemic' or 'Spanish flu' are collected as not preferred terms. The preferred terms are: 'influenza' or 'influenza pandemic, 1918-1919'. Although the controlled vocabularies are clear in their election, most of the literature about '1918 pandemic' is retrievable either by 'Spanish' or by '1918' disjunct, and the dominant word to retrieve literature is 'Spanish' rather than '1918'. This is surprising considering the existence of suitable controlled vocabularies related to health topics, and the modern guidelines of World Health Organization concerning naming of diseases that point out to other preferred terms. A first conclusion is the failure of using controlled vocabularies for a field such as health, and in consequence for WoS and Scopus. This research opens further research questions about which is the role that controlled vocabularies play in the instructions to authors that journals deliver to documents' authors.

Keywords—Controlled vocabularies, indexing, 1918 influenza, information retrieval, keywords, 1918 pandemic, scientific databases.

I. INTRODUCTION

THE role of Knowledge Organization Systems (KOS) and their controlled vocabularies in information retrieval is broadly recognized as relevant feature. Besides, there is a standing demand that editors and databases should consider the effective introduction of controlled vocabularies in their procedures to index scientific literature. On the other hand, the 1918 pandemic is a topic of scientific study one hundred years later, not only by historians but also by researchers from other disciplines. His scientific interest has been fostered by the emergence in the early 21st century of various epidemics of pneumonic diseases caused by viruses, such as COVID-19, which bear similarities to the 1918 pandemic [1]. KOS may also have interest for Library and Information Science, as an example of how terms have a relevant role for indexing and retrieving scientific literature to avoid the documentary noise and silence [2], [3]. In turn, Library and Information Science has a potential to enhance practice of other scientific disciplines or professions, as for instance Health [4]-[8] because an optimal information retrieval through databases should reduce the difficult tasks of manual information retrieval [9] and make easier the clinical information exchange between organizations [10], [11]. Besides, there are works that point out the complexity of subject searching in the catalogs [8], [9] and the dichotomy between keywords from authors versus keywords based on controlled vocabularies [12]. Finally, there is some demand to unify the terminology related to health to make easier the medical history exploitation and research [13]. Considering these features, this paper focuses on controlled vocabularies related to the health field (clinical terminology, electronic health records or research) and their role for storing, classifying, and retrieving relevant literature.

The 1918 flu or influenza pandemic has often been (and still is) referred to as the "Spanish flu" or "Spanish influenza", not just in colloquial language but also in scientific literature. Nowadays, those names are controversial because of reasons regarding the suggested criteria for terminology, the state of scientific knowledge and the information retrieval of related scientific literature: 1) The World Health Organization (WHO) advises against associating names of diseases with names of countries or collectives, to avoid stigmatization of individuals or groups. They even mention "Spanish flu" as an example of name to be avoided for new diseases [14]. 2) Those names equivocally point to a supposed geographical origin of the disease in Spain, although the hypotheses proposed since long

M. Garcia-Alsina is associate professor at the Universitat Oberta de Catalunya, Barcelona, Spain (phone: +34 933263694; fax: 933568822; e-mail: mgarciaals@uoc.edu).

J. Cobarsí-Morales is associate professor at the Universitat Oberta de Catalunya, Barcelona, Spain (phone: +34 933 263 875; fax: 933568822; e-mail: jcobarsi@uoc.edu).

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:16, No:8, 2022

time ago point to geographical origins in other countries [15]; and 3) The controlled vocabularies should achieve consistency in the description of contents and facilitate retrieval, through being properly embedded in databases. In this sense, the need to use different terms to retrieve a single concept such as "1918 pandemic" points out a deficiency on indexing scientific literature in databases. Specifically, it seems (as this paper will show further) there is not a specific term of a controlled vocabulary to describe and index scientific literature referred to 1918 pandemic or if that exists, it has not been properly embedded in scientific databases. So, according to 21st century knowledge and criteria, scientific literature about this topic 'should' be retrievable on a usual basis through "1918 influenza" or "1918 flu" keywords, according to the criteria of avoiding names of countries or collectives. Consequently, in this context, it is relevant to know which terms have been selected by the controlled vocabularies to refer to this "1918 pandemic" and how papers about this pandemic have been indexed and retrieved in practice through scientific databases. In this sense, rather than a thorough discussion of what 'should' be, the present paper sets up as objective to achieve knowledge about what actually 'is'; that could be useful to understand present functioning of scientific literature's index and search, as a basis for future improvements.

It seems that the current functioning of journals and databases may be conditioned by historical inertia and other influences. Hence, this is a relevant case study to obtain interesting insights about needs to improve indexing and retrieving scientific literature in journals and databases.

Considering the previous explanations, the objective of this exploratory research is twofold. Firstly, the research explores which are the terms considered for 1918 influenza in controlled vocabularies for the health field. Secondly, it studies which is the role of controlled vocabularies for storing, classifying, and retrieving. Specifically, the research questions, referred to 1918 influenza related scientific literature, are: 1) Which are the terms preferred in controlled vocabularies for the health field? 2) How papers are indexed in relevant databases such as WoS and Scopus in order to be easily retrieved, considering keywords effective for retrieval vs specific health-controlled vocabularies?

In the remaining of this paper, Section II provides firstly information about what are the most representative controlled vocabularies in the Health field that could be used to index the topic of 1918 pandemic, and secondly, it provides historical background about how 1918 pandemic has been called and what are the hypotheses about geographical origin. Section III presents the methodology used in this research. It follows Section IV that explains the results divided in two parts. Firstly, there are the terms related to this topic in the main controlled vocabularies. Secondly, from a quantitative point of view, it is provided an overview about how scientific papers related to "1918 pandemic" have been classified, which words have been employed, and how these papers may be retrieved through databases in the period 2000-2019. Finally, in Section V, conclusions are set up and future research is suggested.

II. BACKGROUND

It is well known that controlled vocabularies are the result of the process of selecting terms that involves consulting various sources of words and phrases (ANSI/NISO Z39.19-2005) [28]. There is a wide variety because they could be generalist such as Library of Congress Subject Headings (LCSH) (1909) or UNESCO Thesaurus (1977) or specialized for a specific field. For instance, related to Humanities and Social Science the reference is HASSET (Humanities and Social Science Electronic Thesaurus), developed in the 1970s by the UK Data Archive, or related to social science it is the multi-lingual ELSST (European Language Social Science Thesaurus) [29] available in 13 languages, originally based on the monolingual HASSET. ELSST is owned by the Consortium of European Social Science Data Archives (CESSDA) [29]. Related to the health field, several authors have mentioned specific vocabularies, the most relevant seems to be the Medical Subject Heading thesaurus (MeSH) of the National Library of Medicine (NLM) of USA, dated from 1960s [30]. This thesaurus includes the subject heading appearing in MEDLINE (Pub MED), the NLM Catalog and other NLM databases. Nevertheless, some authors point out also the need for more research to build a medical language system to identify clinical terms and index documents, research, and clinical histories [13], [31], [32], [8].

The 1918 pandemic has been so far the largest in human history. Estimates consider between 50 and 100 million deaths [16], [17]. As an emerging phenomenon of unknown origin, the epidemic initially received different colloquial names according to countries, for example: "Flanders Flu" (United Kingdom), "Blitzkatarrh" (Germany), "French flu" (Spain), etc. Later, both in colloquial and in scientific language, this pandemic has been mainly known worldwide by two names: "1918 flu" and "Spanish flu" (flu or influenza).

The name "1918 flu" is obvious. However, the reasons why it is known as the "Spanish flu" are more confusing. The name seems to hint that the first major outbreak of the disease occurred in Spanish territory. This is suggested by analogy with other known epidemics, for example: Russian flu of 1977 and 1889, Hong-Kong flu of 1968, etc. In the former, the name corresponds to the established fact of the country or geographic area where the first massive outbreak occurred. But this is not the case for the 1918 flu. In fact, the current scientific literature has not clearly established the geographical origin of the epidemic. Since early studies on the subject, three geographical options regarding the origin have been considered, all three related with World War I events: 1) North American (and transferred to Europe by the combatants sent to war), 2) European (somewhere on the western front), 3) Asian (and facilitated its diffusion to the United States and Europe by movements of workers or combatants from Asia). These early studies [18]-[20] consider these three possibilities and favor the option of North American origin, having examined documentation and related evidence in detail. For several decades this is the origin accepted as the most plausible, although it has been questioned in recent decades (for instance: [21], [22]). In fact, it is the one that seems to have received the most support over time (for instance, in 21st century papers:

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:16, No:8, 2022

[15], [23]). However, attempts to conclusively resolve this issue in 21st century have been unsuccessful. Anyway, literature review on the subject shows that Spain [15], [18], [20], [21], [23] (a non-belligerent country in the World War I) is not mentioned significantly among the possible geographical origins of the epidemic. Therefore, the widespread adoption of the term "Spanish flu" is surprising in principle. The usual explanation is that the adoption of this term was facilitated by the fact that Spain was a neutral country in the World War I, so that news about the flu circulated without restrictions in the press of this country, unlike in belligerent countries, where war censorship and self-censorship was a general practice. This favored the perception that Spain was the focus and perhaps the origin of the epidemic [24].

Among the pioneers of the use of "Spanish flu, it is worth pointing out the following examples: Swiss Ministry of Health in 1918 [25], the British press as early as May 2018 and the British Medical Journal in August of that same year [26]. The name also began to be used in the United States [27]. Therefore, as early as 1918 this name has been used by official or para-official sources, in English and other Western languages. Anyway, a detailed study of the quick process of adoption for this curious name is not the purpose of our paper, but its trace in nowadays contents in scientific journals and databases. Besides, it is an argument to put in value the controlled vocabularies to index literature in journals and databases. This variety of terms to name the same fact shows how relevant are the controlled vocabularies and its proper application to information retrieval.

III. METHODOLOGY

Considering the first objective mentioned above, to identify which are the suitable terms to index and consequently retrieve literature about 1918 pandemic, two action have been done: a) online registries containing KOS Typologies [33], [34] have been explored to identify controlled vocabularies related to the health field: and b) several controlled vocabularies have been browsed to identify the preferred terms.

The registry selected was "BAsel Register of the Thesauri, Ontologies and Classification" (BARTOC). It has two search options: Basic Search by keywords, and Advanced Search by taxonomy terms. Secondly, controlled vocabularies generalist and specialized related to social science and health have been explored: the LCSH, the UNESCO Thesaurus, HASSET and MeSH. There, it has been identified which are the preferred terms in all these vocabularies.

To approach the second objective, how scientific production about the 1918 pandemic has been indexed and may be retrieved, searches have been conducted in the two most prestigious general scientific databases: WoS and Scopus. Both include journals and databases from several disciplines, including MEDLINE and Pubmed. Searches have been conducted over the period 2000-2019 because during these years the information retrieval may have been improved in a homogenous way rather than in previous decades, and because by 2003 started a renewed interest for 1918 pandemic due to the emergency of SARS epidemic that has increased research about this topic. Four terms were selected considering two criteria and synonyms. That is "country" and "year when the pandemic started", and the synonym. Hence, the following four terms have been selected: "Spanish influenza" and its synonym "Spanish flu", and "1918 influenza" and its synonym "1918 flu". Three search strategies have been applied using these terms (Fig. 1):

- Simple search using a term per search: "Spanish influenza" (SPA-I), "Spanish flu" (SPA-F), "1918 influenza" (1918-I), "1918 flu" (1918-F). This shows which term is the most used. Nevertheless, it fails knowing if authors or journal also use the other term.
- Intersection search between the country name (Spanish) and the year (1918): a) SPA-I OR SPA-F; b) 1918-I OR 1918-F; c) 'both Spanish and 1918' (SPA-I OR SPA-F) AND (1918-I OR 1918-F). This also shows which is the term preferred and if the authors use both.
- Disjunction search between country ("Spanish") and year (1918) excluding the results that content the other term: a) 'Spanish only' (SPA-I OR SPA-F) AND NOT (1918-I OR 1918-F); b) '1918 only' (1918-I OR 1918-F) AND NOT (SPA-I OR SPA-F). This confirms us which is the term more used, and which ones produce documentary silence or noise.

Simple search	Spanish influenza (SPA-I) Spanish flu (SPA-F) 1918 influenza (1918-I) 1918 flu (1918-F)							
Intersection between Spanish and 1918	Spanish influenza	OR	Spanish flu					
	1918 Influenza	OR	1918 Flu					
	"both Spanish and 1918"	Spanish Influenza	OR	Spanish flu	AND	1918 influenza	OR	1918 flu
disjunction between Spanish and 1918	"Spanish only"	Spanish influenza	OR	Spanish flu	AND NOT	1918 influenza	OR	1918 flu
	"1918 only"	1918 influenza	OR	1918 flu	AND NOT	Spanish influenza	OR	Spanish flu

Fig. 1 Search strategy

IV. RESULTS

Firstly, this section, presents which are the most representative controlled vocabularies related to the health field and their preferred terms to classify and index the scientific literature about "1918 pandemic". Secondly, it exposes which have been the terms used by databases to index the scientific production and how it should be retrieved.

A. Controlled Vocabularies and Preferred Terms

Searches in BARTOC show that term "epidemic" or

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:16, No:8, 2022

"pandemic" do not retrieve controlled vocabularies. Nevertheless, in BARTOC "Grippe" and "influenza" are considered. The term "Grippe" retrieves the WO2 Thesaurus for the Second World War. The term "influenza" retrieves some vocabularies from the medical field, from which more searches about influenza could be done, as for instance the vocabulary of "Disease Ontology" (Fig. 2). But the "1918 pandemic" is not included in BARTOC.

A part of BARTOC, the most relevant vocabularies related to the health field, has been explored. The LCSH considers 30 labels for Influenza (Fig. 3). This research is focused on three terms: a) "influenza" and their variants "Flu", "Flu respiratory", "Grippe" and "Respiratory Flu"; b) "Influenza Epidemic, 1918-1919"; and "Influenza Epidemic, 1918-1919", in literature.
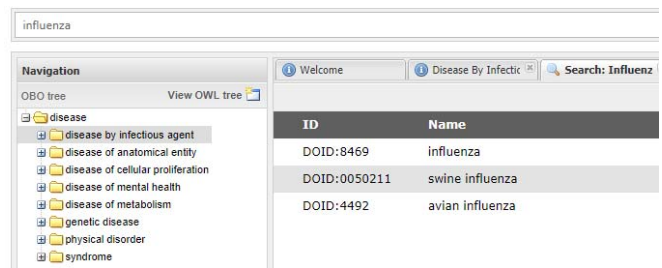

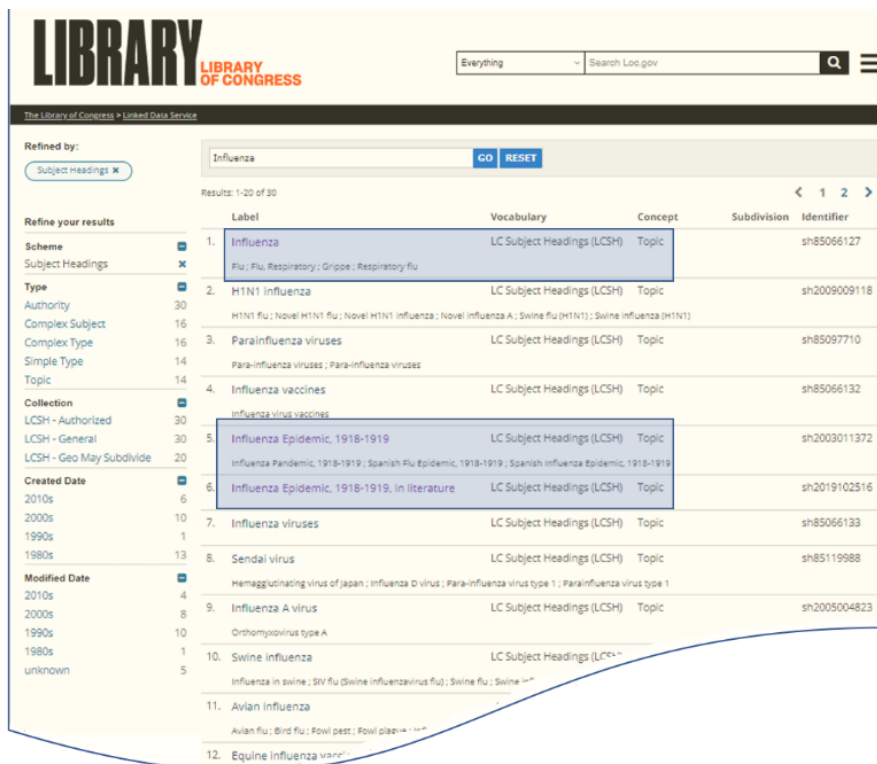Fig. 2 Influenza in Disease Ontology

Fig. 3 Labels for Influenza in LCSH

"Influenza Epidemic 1918-1919" points out as variants "Influenza Pandemic", "Spanish Flu" and "Spanish influenza Epidemic". These variants are not collected in the other labels. That reflects the use of other terms to index scientific literature about *1918 pandemic*, but the preferred term in this vocabulary is the *year* and not the *country name* (Fig. 4).

The Medical Subject Heading (MeSH) collects the term "Spanish Flu Pandemic 1918-1919", although the preferred term is "Influenza Pandemic 1918-1919" (Fig. 5). That is relevant for this research as this vocabulary is available since the 1960s.

Finally, HASSET considers "Influenza" in the context of "diseases" and "respiratory infections". It is the preferred term in front of the term "Flu". Variants such as "Spanish Influenza" or "1918 Influenza" are not considered (Fig. 6).

B. Retrieval of Scientific Papers about 1918 Pandemic

Here the results are presented about which terms and search strategy are the most suitable for retrieving related scientific literature. It follows the following structure: 1) simple search using only one term; 2) search using the intersection of the two kinds of terms (country and year); 3) searching with one term and excluding the other term into the results; and 4) the comparative between intersection and disjunction.

When the search is conducted with only one term, some divergencies are appreciated between the most suitable term in WoS and Scopus. The most relevant term in WoS is the idea of "Year" ("1918 Influence") and consequently it avoids the use of the name of a country or a collective name. On the contrary, in Scopus the most relevant term is the country name ("Spanish flu"). Nevertheless, the addition of "Spanish influenza" and

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:16, No:8, 2022

"Spanish flu" in WoS is more than 50% (57,19%). That means that apparently the name of a country is also relevant to retrieve literature about the topic of pandemic in WoS. This tendency is stronger in Scopus where the addition of the two terms that contents the name of a country is the 68,58%. The reason could be that some papers have been indexed with both terms. Hence, it is necessary more search to conjunct and disjunct the different terms, to explore which term is more relevant (Fig. 7).
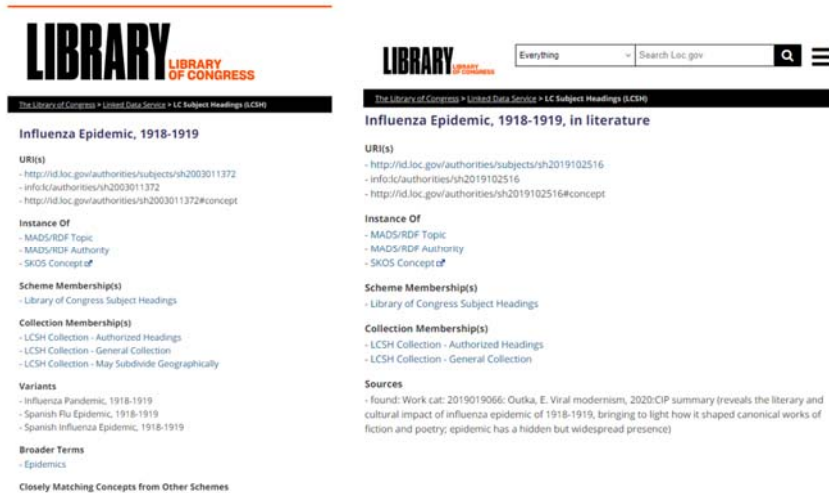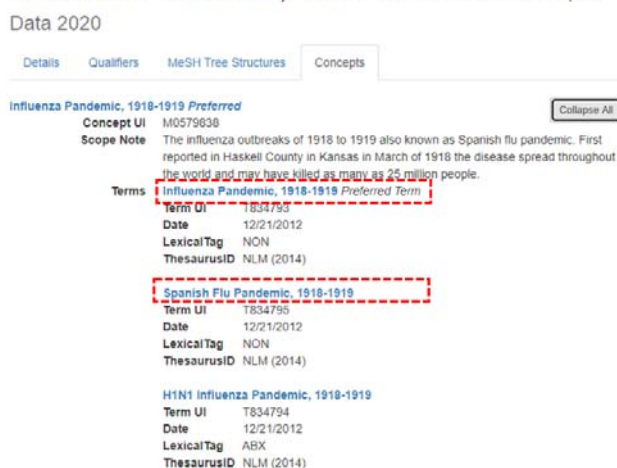

Fig. 4 Influenza Epidemic in LCSH

Fig. 5 Search in the Medical Subject Heading (MeSH)

According to the second search strategy, when we combine the terms, the findings confirm the previous results. So, the country term has the most weight, and this is stronger in Scopus than in WoS. There are few papers that are retrieved both by the country name and the year name together. Consequently, we can conclude that the term referred to a country ("Spanish") is most relevant that the term referred to the year (1918) when the pandemic happened (Fig. 8).

Thirdly, the search using only one word supports the result that the most used term is "Spanish Influence" or "Spanish flu". The search excluding one of the other terms (*country* or *year*) also shows preference for the use of the country name. WoS shows more balance between the use of the two terms than Scopus. Both WoS and SCOPUS support the view that to avoid documentary silence both terms (country and year) should be used together (Fig. 9).
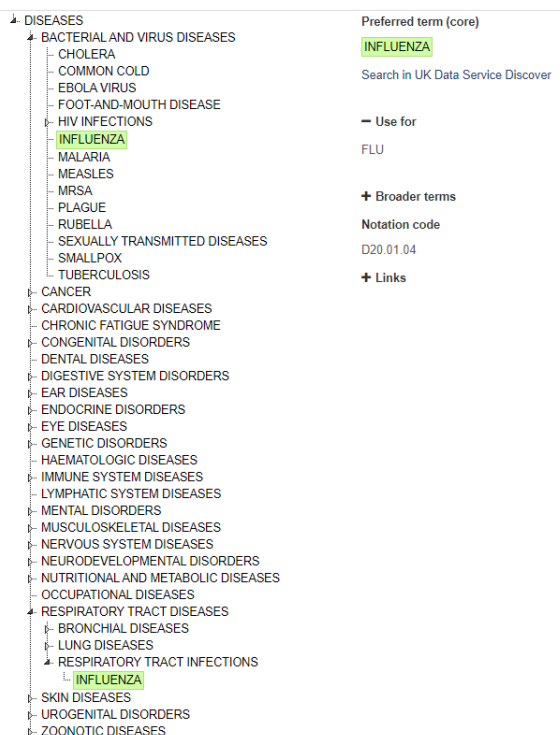

Fig. 6 Influenza in HASSET

Finally, Fig. 10 shows as a synthesis the results of the searches through a Venn's diagram visualization. Graphic shows similar results for both databases: First, although country and year are equivalent keywords for the same concept, the result of the search is highly dependent on which keyword is chosen, intersection is very small (9% of total results in WoS,

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:16, No:8, 2022

13% of total results in Scopus). Secondly, country keyword is dominant in both databases, in a more marked way in Scopus (51% of total results in WoS, 62% of total results in SCOPUS).

ISI WoS				Scopus			
SPA-I	SPA-F	1918-I	1918-F	SPA-I	SPA-F	1918-I	1918-F
333	343	450	56	630	324	387	50
28,17%	29,02%	38,07%	4.74%	45,29%	23,29%	27,82%	3,6%
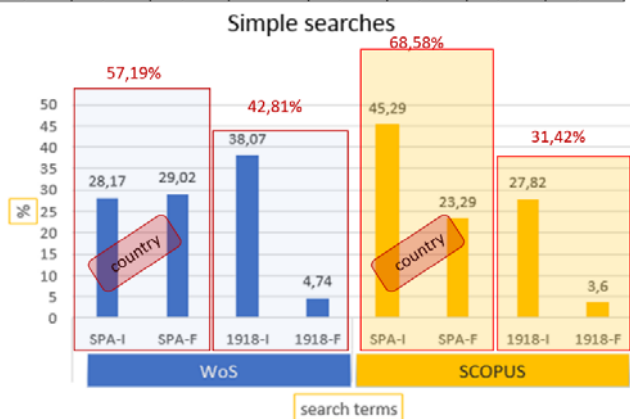
Fig. 7 Results of simple search, using one word

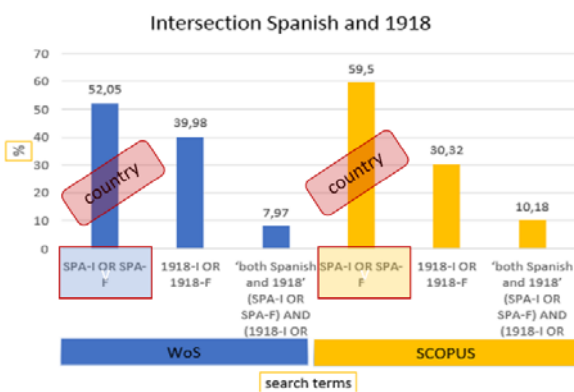| ISI WoS			Scopus		
SPA-I OR SPA-F	1918-I OR 1918-F	'both Spanish and 1918' (SPA-I OR SPA-F) AND (1918-I OR 1918-F)	SPA-I OR SPA-F	1918-I OR 1918-F	'both Spanish and 1918' (SPA-I OR SPA-F) AND (1918-I OR 1918-F)
647	497	99	848	432	145
52,05%	39,98%	7,97%	59,5%	30,32%	10,18%

Fig. 8 Intersection between "Spanish" and "1918"

V. CONCLUSIONS

A first conclusion is the failure of the application of controlled vocabularies for a field such as Health, and in consequence for WoS and Scopus. Both terms "Spanish" and "1918" are used to index the scientific production, although the term more used is that related to the country. Hence, there is a relevant disjunction between what is retrieved with "Spanish" or what is retrieved with "1918". Consequently, to avoid the documentary silence, searches should be done both with Spanish and 1918 terms. In the same sense, that is how *1918 pandemic* related scientific papers may be retrieved through databases. This research has detected that search through both

"Spanish" and "1918" retrieves just 9% of results in WoS and 13% of results in SCOPUS. This is a hint of deficiencies in indexation practices by authors and/or journals' editors and/or databases, which deserve further study. According with this context, there are controlled vocabularies to describe and index scientific literature referred to *1918 pandemic* and produced since 2000. Nevertheless, it seems that they have not been considered effectively by authors and/or journals and/or databases. Consequently, this is a hint for journals' editors and databases to consider the effective introduction of controlled vocabularies in their procedures to index scientific literature. If indexing scientific literature considers controlled vocabularies, search for literature reviews would be easier and more effective.

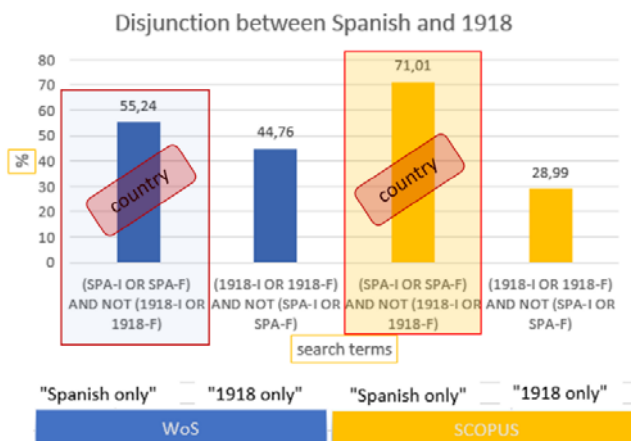| ISI WoS		Scopus	
'Spanish only' (SPA-I OR SPA-F) AND NOT (1918-I OR 1918-F)	'1918 only' (1918-I OR 1918-F) AND NOT (SPA-I OR SPA-F)	'Spanish only' (SPA-I OR SPA-F) AND NOT (1918-I OR 1918-F)	'1918 only' (1918-I OR 1918-F) AND NOT (SPA-I OR SPA-F)
606	491	703	287
55,24%	44,75%	71,01%	28,99%

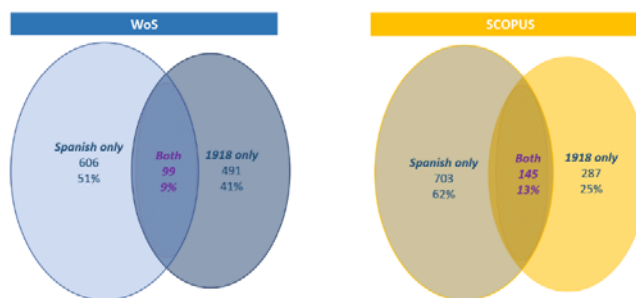Fig. 9 Disjunction between Spanish and 1918

Fig. 10 Results using both terms

These first results suggest that more research is needed about practices, guidelines, and procedures, for authors, journals' editors, and scientific databases managers, regarding indexation and use of keywords and terminology. This research opens other research question about which is the role that controlled vocabularies play in the instructions to authors that journals deliver to documents' authors. This topic will be a part of our future research and this is relevant beyond the specific context

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:16, No:8, 2022

of this example. On the other hand, more specific of this example, it could be worth specific further research about the very quick process and long-term adoption in a scientific context of a country related name for a disease.

REFERENCES

[1] Liu, W. J., Bi, Y., Wang, D., Gao, G. F. On the Centenary of the Spanish Flu: Being Prepared for the next Pandemic. Virologica Sinica 33, 463-466 (2018).

[2] Bouramoul, A., Kholladi, M.K., Doan, B.L. How Ontology Can be Used to Improve Semantic Information Retrieval: The AnimSe Finder Tool. International Journal of Computer Applications, Foundation of Computer Science 21 (9), 48-54 (2011).

[3] Alvite Díez, M.L. Tendencia en la investigación sobre recuperación de información jurídica. Revista Española de Documentación Científica 26(2), 191-212 (2003).

[4] Boukhari, K., Omri, M.N. DL-VSM based document indexing approach for information retrieval. Journal of Ambient Intelligence and Humanized Computing 1-12. https://doi.org/10.1007/s12652-020-01684-x (2020).

[5] Utechta, J., Ballb, J., Bowmanc, S. M., Doddb, J., Judkinsd, J., Maxsone, R. T., Nabaweesif, R., Pradhanc, R., Sanddalb, N. D., Winchellg, R. J., Brochhausena, M. Development and Validation of a Controlled Vocabulary: An OWL Representation of Organizational Structures of Trauma Centers and Trauma Systems. In Ohno-Machado, L. and Séroussi, B. (eds.) MEDINFO 2019: Health and Wellbeing e-Networks for All (2019)

[6] Chae, J., Cho, I., Yoo, A., Kim, Y. Analysis of a Locally Controlled Vocabulary in an Electronic Health Records for Evidence-Based Inpatient Fall-Prevention Care. Nursing Informatics 78 (2018).

[7] Soualmia, L. F., Darmoni, S. J. Combining different standards and different approaches for health information retrieval in a quality-controlled gateway. Informational Journal of Medical Informatics 74(2-4), 141-150 (2005).

[8] Darmoni, S., Soualmia, L. F., Letord, C., Jaulent, M.C., Griffon, N., Thirion, B., Névéol, A. Improving information retrieval using Medical Subject Headings Concepts: a test case on rare and chronic diseases. Journal of the Medical Library Association, 100(3):176-183. DOI: 10.3163/1536-5050.100.3.007. PMID: 22879806; PMCID: PMC3411256. (2012).

[9] Lelong, R., Cabot, C., Soualmia, L. F. Semantic Search Engine to Query into Electronic Health Records with a Multiple-Layer Query Language. SGIR 16 Juli, 2016 Pisa, Italy. ACM (2016).

[10] Kierkegaard, P., Kaushal, R., Vest, J. R. Information Retrieval Pathways for Health Information Exchange in Multiple Care Settings. The American journal of managed care 20(11 Spec No. 17), SP494-501 (2014).

[11] Chae, J.; Cho, I.; Yoo, A.; Kim, Y. Analysis of a Locally Controlled Vocabulary in an Electronic Health Records for Evidence-Based Inpatient Fall-Prevention Care. Nursing Informatics (2018)

[12] Gross, T.; Taylor, A. G.; Joudrey, D. N. Still a Lot to Lose: The Role of Controlled Vocabulary in Keyword Searching. Cataloging & Classification Quarterly, 53:1, 1-39 (2015).

[13] Bousquet, C; Souvignet, J; Sadou É; Jaulent M-C; Declerck G. Ontological and Non-Ontological Resources for Associating Medical Dictionary for Regulatory Activities Terms to SNOMED Clinical Terms with Semantic Properties. Frontiers in Pharmacology (2019) 10;10:975. doi: 10.3389/fphar.2019.00975. PMID: 31551780; PMCID: PMC6747929. (2019)

[14] World Health Organization. World Health Organization Best Practices for the Naming of New Infectious Diseases. May 2015. https://www.who.int/topics/infectious_diseases/naming-new-diseases/en/ last accessed 2020/10/9

[15] Barry, J. M. The site of origin of the 1918 influenza pandemic and its public health implications. Journal of Translational Medicine, 2:3 (2004).

[16] Knobler S, Mack A, Mahmoud A, Lemon S, (eds.) "1: The Story of Influenza". The Threat of Pandemic Influenza: Are We Ready? Workshop Summary (2005). Washington, DC: The National Academies Press. pp. 60–61 (2005).

[17] Johnson, N.P., Mueller, J. Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic. Bulletin of the History of Medicine. 76 (1), 105–15. doi:10.1353/bhm.2002.0022. PMID 11875246 (2002).

[18] Jordan, E. Epidemic influenza First edition. Chicago: AMA; (1927).

[19] Thomson D., Thomson R: Influenza. Annals of the Pickett-Thomson Research Laboratory First edition. Baltimore: Williams and Wilkens; (1934).

[20] Burnet, F.M., Clark, E. Influenza: a survey of the last fifty years Melbourne.: Macmillan Co; (1942).

[21] Oxford, J.S., Lambkin, R., Sefton, A., Daniels, R., Elliot, A., Brown, R., Gill, D. A hypothesis: the conjunction of soldiers, gas, pigs, ducks, geese and horses in northern France during the Great War provided the conditions for the emergence of the ''Spanish'' influenza pandemic of 1918–1919. Vaccine 23, 940–945 (2005).

[22] Shortridge, K.F. The 1918 'Spanish' flu: pearls from swine? Nature Medicine, vol 5, number 4, April 1999, 384-385 (1999).

[23] Taubenberger, J.K., Morens, D.M. 1918 influenza: the mother of all pandemics. Emerg Infect Dis 2006; 12, 15–22 (2006).

[24] Rosenheck, J. The so-called "Spanish" flu. Doctor's Review. http://www.doctorsreview.com/history/nov05-history/ (2005) last accessed 2020/10/09

[25] Mecking, E. Spaanse griep en Eerste Wereldoorlog Het drama van 1918. De Eerste Wereldoorlog 1914 – 1918. https://www.wereldoorlog1418.nl/spaanse%20griep/index.html last accessed (2006) 2020/10/09

[26] Bennett, H. Spanish Flu and the history of pandemic propaganda. Prospect, March 19 2020. https://www.prospectmagazine.co.uk/politics/coronavirus-spanish-flu-1918-pandemic-from-spain (2020) last accessed 2020/10/09

[27] Soper, G. A. The Influenza Pneumonia Pandemic in the American Army Camps during September and October, 1918. Science, Nov. 8, 1918, New Series, Vol. 48, No. 1245 (Nov. 8, 1918), pp. 451-456 (1918). https://www.jstor.org/stable/1641514?seq=1#metadata_info_tab_contents last accessed 2020/10/09

[28] ANSI/NISO Z39.19-2005. Guidelines for the construction, Format, and Management of Monolingual Controlled Vocabularies. (2005).

[29] UK Data Service. About ELSST. https://elsst.ukdataservice.ac.uk/ last accessed 2020/10/09

[30] Rogers, F B (Jan 1963). "Medical subject headings". Bulletin of the Medical Library Association, 51(1):114–116 (1963).

[31] Soualmia, L.F., Sakji, S., Letord, C., Rollin, L., Massari, P., Darmoni, S.J. Improving information retrieval with multiple health terminologies in a quality-controlled gateway. Health Information Science Systems, 4(1):8. DOI: 10.1186/2047-2501-1-8. PMID: 25825660; PMCID: PMC4341235 (2013).

[32] Rosenbloom, S.T., Miller, R.A., Johnson, K.B., Elkin, P.L., Brown, S.H. Interface terminologies: facilitating direct entry of clinical data into electronic health record systems. Journal of American Medical Information Association13(3), 277-88. doi: 10.1197/jamia.M1957. Epub 2006 Feb 24. PMID: 16501181; PMCID: PMC1513664. (2006)

[33] Bratková, E.; Kucerová, H. Knowledge Organization Systems and Their Typology. Revue of Librarianship. 25(2):1-25.

[34] Park, Z.; Gnoli, C; Morelli, D. The Second Edition of the Integrative Levels Classification: Evolution of a KOS, 5(1):39-50.

Montserrat Garcia Alsina (Barcelona, Spain, 1961) holds a bachelor in Geography and History (Universitat de Barcelona, Barcelona, Spain, 1984), a bachelor in Information and Documentation (Universitat Oberta de Catalunya, Barcelona, Spain, 2002) and a Ph.D. in Knowledge and Information Society (Universitat Oberta de Catalunya, Barcelona, Spain, 2011).

She has been information and records manager in an engineering company (Barcelona, Spain, 1991-2007). She is currently Associate Professor of Information and Communications Studies at Universitat Oberta de Catalunya, Barcelona, Spain). She has published a book about bog data (2017), and two more books about competitive intelligence (2009, 2012). She has published in indexed journals such as ASLIB Information Proceedings, Revista Española de Documentación Científica and El Profesional de la Información. Her main research interests are about information and knowledge management, and competitive and territorial intelligence.

Dr. Garcia-Alsina takes part of the ISO/TC 46 for Information and Documentation and the Spanish UNE CTN50 for Documentation.

Josep Cobarsí (Girona, Spain, 1967) holds a bachelor in telecommunications engineering (Universitat Politècnica de Catalunya, Barcelona, Spain, 1994) and a Ph.D. in organization management (Universitat de Girona, Spain, 2005).

He has been Webmaster and IT librarian at Universitat de Girona (1995-2000). He is currently Associate Professor of Computer Science, Multimedia and Telecommunication Studies at Universitat Oberta de Catalunya (Barcelona, Spain). He has published book about information disasters (Pequeños y grandes desastres de la información, Barcelona: Infonomía,) and about information systems (Los sistemas de información en la empresa, Barcelona: UOC, 2011). He has published in indexed journals such as ASLIB Information Proceedings, Revista Española de Documentación Científica and El Profesional de la Información. His main research interests are about knowledge management, digital behavior and disaster management.

Dr. Cobarsí is affiliated to Association for Information Science and Technology (ASIS&T), being Spain country representative of this association.