

**Ann M. Graf** — School of Information Studies, University of Wisconsin-Milwaukee  
**Richard P. Smiraglia** — School of Information Studies, University of Wisconsin-Milwaukee

## **Transition in Education: Domain Analysis from the *Encyclopedia of Milwaukee***

### **Abstract**

Encyclopedism and its products, encyclopedias are an important form of knowledge organization system. Encyclopedias are deliberate cultural syntheses. Local online encyclopedias represent an emerging trend in the development of KOSs. The Encyclopedia of Milwaukee is one of the most recent additions to this trend. An interesting problem for research is how to visualize the domain of an evolving online encyclopedia. In this study, the base bibliography for one disciplinary cluster—education—is submitted to bibliometric domain analytical techniques. While parameters of a typical domain emerge, other atypical results also emerge, raising questions about the differences represented in encyclopedia domains as well as questions about the historical evolution of scholarship. Nonetheless, transitions in the evolution of the domain of education in Milwaukee, including increased author productivity and enhanced granularity, as well as local cultural distinction, clearly emerge.

### **Introduction**

Encyclopedism, the notion that all knowledge can be brought together in one place, and the encyclopedic products that emerge from it, is a form of knowledge organization system (KOS) in the purest sense. Encyclopedias are important KOSs because they reflect accumulated scholarship and illuminate the pathways between diverse bits of knowledge, embodying what Patrick Wilson calls exploitative bibliographic power (1968). “The encyclopedia format draws its strength from the scholarship produced in other research formats, including local histories, dissertations, scholarly monographs, and interpretive syntheses. However, by creating the structure to integrate and elaborate the connections, debates, and controversies inherent in its source material, the encyclopedia plays a major role in the systematic development of additional knowledge” (Seligman and Anderson 2012, 5). Encyclopedias are not mere repositories of stored knowledge, they also are deliberate cultural syntheses.

Local online encyclopedias represent an emerging trend in the development of KOSs, drawing together historians, educators, students of all ages and those with a stake in their own community. The Encyclopedia of Milwaukee (EMKE) is one of the most recent additions to this trend. The EMKE project will result in the publication of a traditional monograph, tentatively scheduled for release in 2016, as well as an interactive website, which not only will delve deeper into the historical and present scholarship used by researchers to develop the print volume, but will provide a platform for community input—it will extend the encyclopedic model by attenuating two-way cultural synergy. It is hoped that visitors to the site will add to the knowledge already presented by contributing their personal stories and images of the city.

In knowledge organization, domain analysis is a set of research techniques that can be applied to understand the intension, extension, and underlying ontology of a domain. According to Smiraglia (2012, 114) a domain is: "a group with an ontological base that reveals an underlying teleology, a set of common hypotheses, epistemological consensus on methodological approaches, and social semantics." The historians,

bibliographers, authors, etc. contributing to the EMKE represent a domain in the process of developing an important KOS. Many domain-analytical techniques are bibliometric (Hjørland and Albrechtsen 1995), and bibliometric techniques have been used to good effect in information studies to understand processes of scientific communication (De Bellis 2009). In knowledge organization (Smiraglia 2012), domain analytical studies have been used primarily to understand the ontology within domains, for the purpose of developing KOSs. However, most studies have analyzed formal scholarly disciplines, such as music information retrieval, or knowledge organization itself. There has been no prior domain analysis of the working scholarship of an encyclopedia. And, in particular, there is no prior research about the extension or intension of the various domains that together comprise the components of an online encyclopedia. Our immediate access to the workings of an evolving online encyclopedia have provided us with a unique opportunity to analyze it as a distinct domain.

The Encyclopedia of Milwaukee project has provided insight into the history of the city and the progression of related scholarship over the course of the last century. The bibliography was created by one of the authors of this paper to serve as a starting point for research for those who would be enlisted to write individual entries for the encyclopedia. The bibliography therefore uses categories that mirror those of the encyclopedia (see figure 1); these categories represent an emergent taxonomy, which we have described elsewhere (Graf and Smiraglia 2012). This large bibliography currently has over 2000 resources. Fourteen categories resulted from the monthly editorial board meetings for the encyclopedia (Education, Religion, Places, etc.) and it is on one specific and easily identifiable subset of the bibliography that this preliminary study will focus.

**Figure 1.** List of categories developed for the EMKE

<b>Arts and Culture (AC)</b>	Museums, dance, theater, libraries, music, foodways, visual art
<b>Business, Industry, Labor and Agriculture (BILA)</b>	Companies, industries, trade associations, unions
<b>Education (E)</b>	Universities, public schools, private schools, trade schools, kindergartens, funding, chapter 220?, school choice, freedom schools, German-American academy, medium of instruction debate/Bennett law, Indian Community School
<b>Gender and Sexuality (GS)</b>	suffrage movement, women's liberation, GLBT, prostitution, women's networking and organizations, women religious
<b>Health and Medicine (HM)</b>	Hospitals, health networks, epidemics, sanitation, public health, pharmaceuticals
<b>Media and Communications (MC)</b>	Film, radio, television, print media, magazines, newspapers
<b>Natural and Built Environment (NBE)</b>	Architecture, geology, water, rivers, the lake, topography, weather, climate, parks (the lake means Lake Michigan)
<b>Places: Neighborhoods and Suburbs (P)</b>	Annexation, grid, street naming and street numbering, suburbs, some version of the neighborhoods
<b>Politics, Law and Government (PLG)</b>	Public order and major conflicts, riots, civil rights demonstrations, strikes, socialism, city government, county government, red light district/prostitution/vice, County Grounds
<b>Race and Ethnicity (RE)</b>	Groups of people, defined with minimal numerical threshold (approximately 500 to 1000 minimum) or evidence of institutional presence
<b>Religion (R)</b>	Lumping rather than splitting – Roman Catholicism, Orthodox, Protestantism, Judaism, Islam, Hinduism, Buddhism, Sikhs, etc.
<b>Sports and Leisure (SL)</b>	Festivals, biergartens, gambling, parks, specific stadia (Miller Park, County Stadium, Mecca, Borchert Field, Pettit Center), teams (pro and amateur), kinds of sports (bowling, soccer, basketball, baseball, cycling, sailing, bocce, sheephead/card games)
<b>Third Sector and Organizations (TS)</b>	Social services, charity, philanthropy, particular foundations
<b>Transportation, Infrastructure, and Public Services (TIPS)</b>	Airports, public surface transit, trolley, rail, interurbans, busses, plank roads, streets, boulevards, tollways, interstate highways, bicycling, canals, river and harbor, sewers and water, electric, gas, telecommunications, waste disposal

## Methodology

An interesting problem for research is to discover how we can visualize the emergence of the domain that is producing the EMKE. As a preliminary study, we have chosen to explore the transition of the domain of education in Milwaukee from its early appearance in the literature of the late 1800s to the present using bibliometric methods. In other words, we have chosen one particularly bibliographically coherent segment of the encyclopedia as an entry point for the analysis of the domain of an online encyclopedia. Specifically, domain analysis is used to visualize the intension and extension of the domain—in this case we examine the supporting bibliography for the encyclopedic treatment of “education in Milwaukee.” Our domain analysis is undertaken primarily through citation analysis and co-word analysis. We hope to observe the ways in which an encyclopedic domain based on a stable bibliographic stratum is like or differs from a community of academic scholars. The analysis reported in this paper, then, is limited to the domain-specific characteristics derived from the base bibliography for “education in Milwaukee,” which contained 177 citations. Therefore specific limitations of this study are first that we have limited ourselves to only one of fourteen disciplinary-based segments of the encyclopedia, and second that

we are analyzing the base bibliography alone, and therefore neither the domain cited within nor the production of the encyclopedia's authors.

### Education in the EMKE

Table 1 shows the distribution of the works in the bibliography for education sorted by type of publication.

**Table 1.** Publication types by number of occurrence

Type of publication	Frequency
Journal article	59
Monograph	48
PhD dissertation	26
Masters thesis	24
Report, commissioned	12
Chapter of monograph	5
Unpublished paper	2
Working paper	1
Total	177

For the present domain analysis, newspaper articles were removed from the study. More research into relevant local newspapers is intended for future study, but was outside the scope of the present study.

The 59 journal articles in Table 1 appeared in 41 separate journals. However, only one journal contained more than three items, *The Wisconsin Magazine of History* with 12 items. This was followed by *The Journal of Negro Education* with 3 items, four journals with two items each, and 35 journals containing only one item each. *The Wisconsin Magazine of History* focuses on general history in Wisconsin, while *The Journal of Negro Education*, published by Howard University in Washington, D.C., has a specific focus on the education of black people (<http://www.journalnegroed.org/>) but no specific focus on Wisconsin.

The 177 total items listed in the bibliography relating to education cite 152 individual authors. Of these, only one author wrote four items, four authors wrote three items apiece, and 12 authors each wrote two items. The remaining 112 authors were each cited only once. The upper portion of this distribution can be seen in Table 2.

**Table 2.** Author productivity frequencies

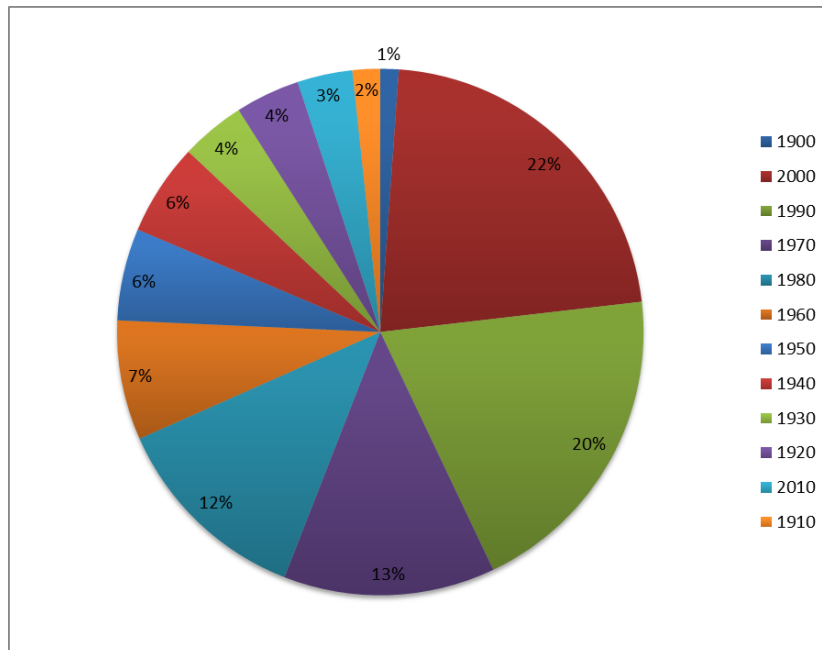
Author	Frequency
Levine, Marc V.	4
Klotsche, J. Martin	3
Fernández, Ricardo R.	3
Goldberg, Bettina	3
McGroarty, Daniel	3
Baez, Luis A. (Tony)	2
Baruch, Steven	2
Birnbaum, Robert	2
Dahlk, William (Bill)	2

Fritz, Angela	2
Kellogg, Louise P. (Phelps)	2
Kieckhefer, Grace Norton	2
Montgomery, Theodore Van Tifflin	2
Nelsen, James Kenneth	2
Pedroni, Thomas (C.)	2
Peterson, Walter F.	2
Rast, Joel	2

The ubiquitous Bradford-like distribution with the long tail of single-author citations is certainly present, but unlike domain analyses of scholarly domains there is no apparent research front represented by a small number of highly productive authors. Rather, the distribution, even at the top, is relatively flat. Seventeen authors produced 40 of the 177 works, and only 5 authors at the very top produced 16 works. It is not immediately apparent from this analysis why this is the case. Several possibilities present themselves. For instance, it could be a function of the bibliography's role in the EMKE project, or perhaps a result of the deliberate selection of works into the bibliography for the encyclopedia. Or it could reflect a lower-productivity but more highly populated domain for education.

We next arrayed the citations chronologically. The 177 works ranged in date of publication from 1901 to 2013. A frequency distribution by decade is depicted in Figure 2.

**Figure 2.** Chronological distribution



In fact, the works cover the entire chronological range with roughly equal frequencies in the first seven decades of the twentieth century, but the majority of works (67%) were produced after 1970. This result raises questions as well. Is this the natural distribution of research over the twentieth century in general, or is it somehow demonstrably representative of the domain of education? To see whether we could observe change over time, we counted publication types from the first three decades, the midpoint (1950-1969) and the last two decades of the distribution. This is visualized in Table 3.

**Table 3.** Change in distribution of publication type

Period	Total # citations	monograph	report	journal	magazine	thesis	dissertation
1900-1920	5	40%	20%	20%	20%	0%	0%
1950-1960	11	27%	9%	0%	27%	37%	0%
2000-2013	45	27%	9%	44%	2%	7%	11%

Just this brief visualization shows one chronological transition in the domain over time. The number of works increases fourfold, and the predominant publication type shifts from monograph to journal article. The relative contribution of reports remains the same over time. The influence of the thesis falls as that of the dissertation rises. As we noted before, this transition cannot be fully explained—it could reflect the shifting dynamics in scholarship in general over the twentieth century, or it could reflect the influence of scientific methods in education toward the end of the time period.

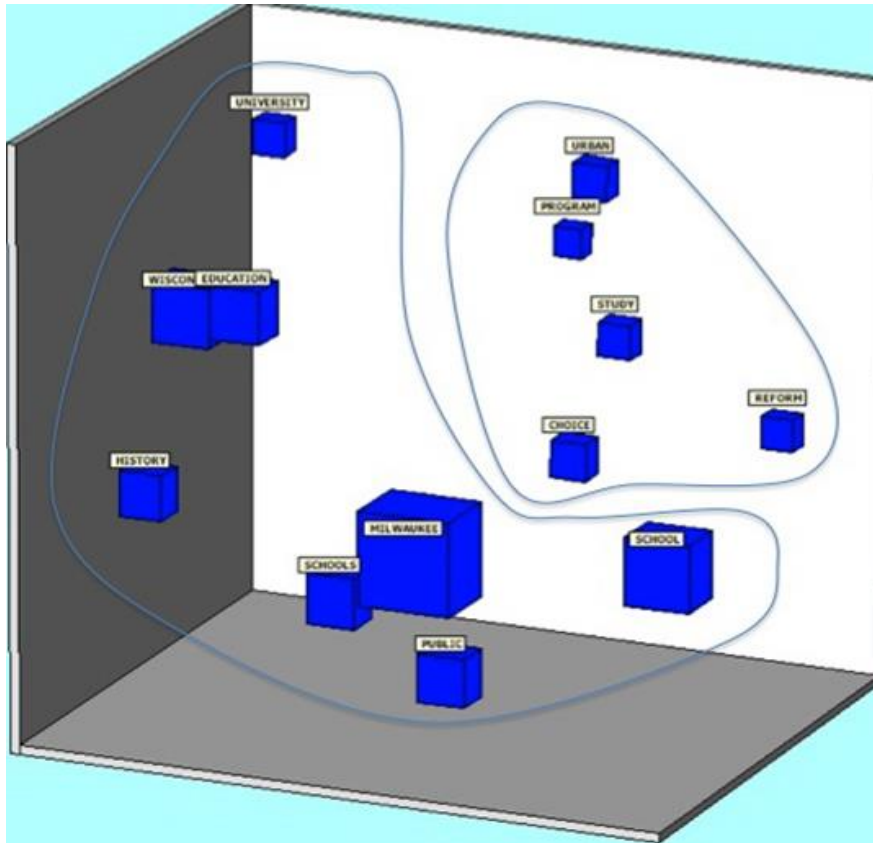
A similar look at author productivity at the same three points in time also shows a transition. In the period 1900-1929 only one of 10 authors (Louise P. Kellogg) produced 2 works. In the period 1950-1969 again only one author among 21 (Grace Norton Kieckhefer) produced 2 works. But in the final two decades from 2000-2013 four authors among 39 produced multiple works (Marc V. Levine 4; James Nielsen, Thomas Pedroni and Joel Rast 2 each). The same interpretive caveats as in the preceding paragraph apply here. A transition is observable, but causality cannot be attributed based on the data in this study alone.

Finally, we analyzed keywords across the entire distribution as a means of understanding the intension of the domain. Table 4 shows the most frequently used keywords extracted using WordStat™ software.

**Table 4.** Title keywords by number of occurrence

Milwaukee	99
School (Schools)	78
Wisconsin	40
Education	31
Public	26
History	22
Choice	15
Urban	14
Educational	13
Study	12

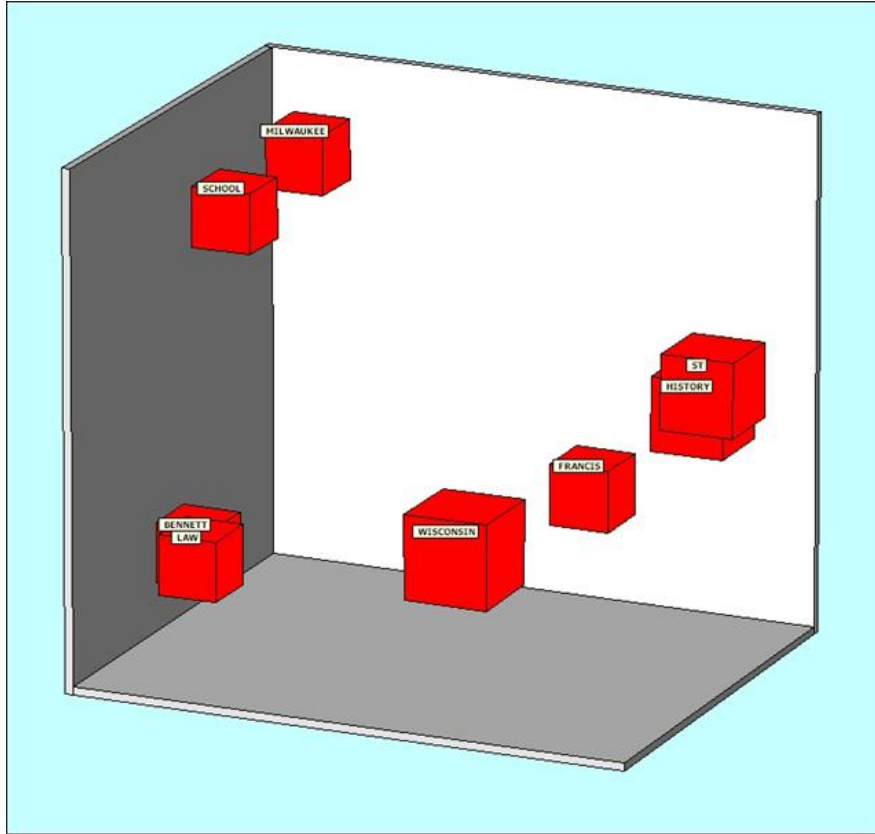




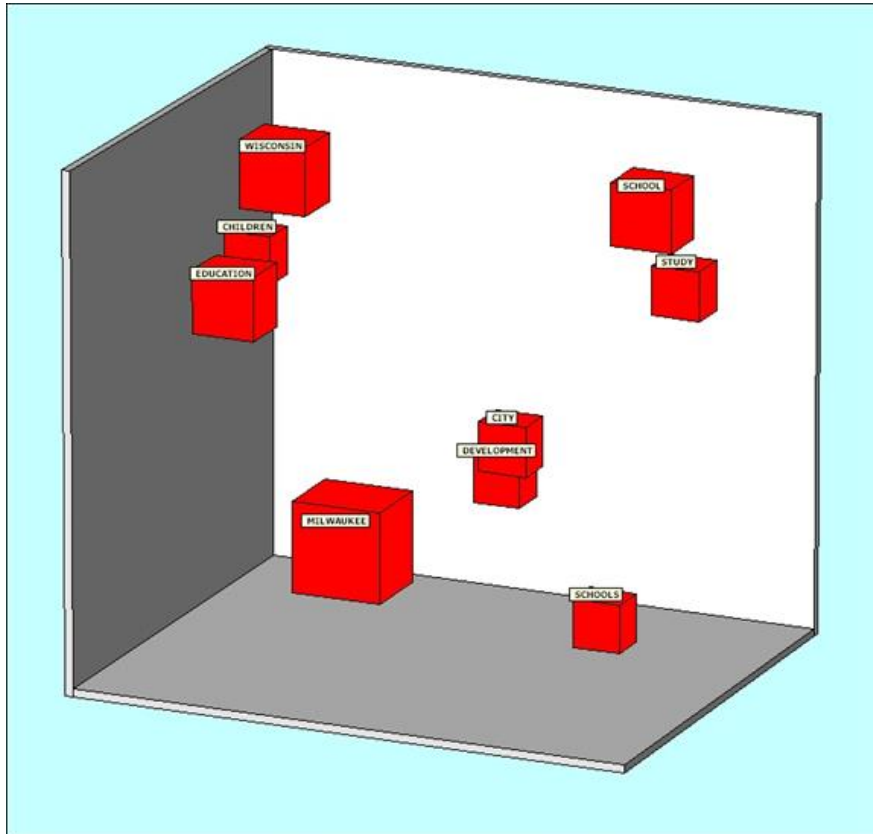
The goodness of fit statistics tell us this is a good model of the data. The most prominent term is “Milwaukee” and the cluster on the lower right shows an emphasis on Milwaukee public education. The cluster on the upper right shows the leading terms such as “reform” and “choice” that represent the granularity apparent in the entire distribution. To visualize thematic change over time we ran WordStat™ to produce MDS plots from the three chronological points of view used above. These plots are arrayed in Figures 5-7.

**Figure 5.** Co-word analysis of title keywords 1900-1929

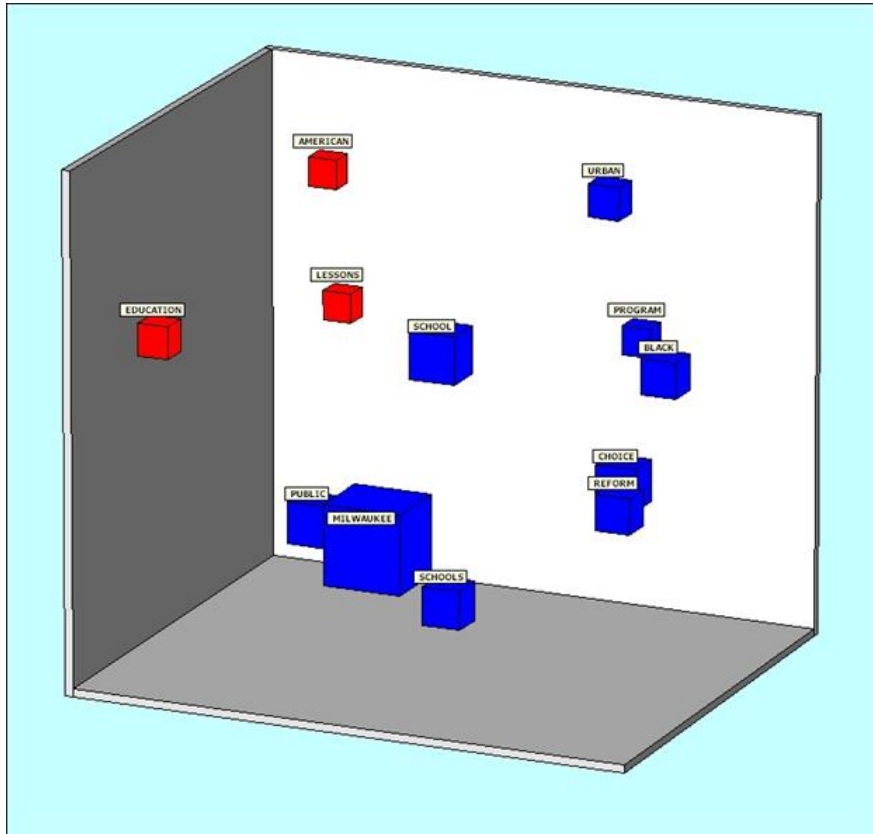




**Figure 6.** Co-word analysis of title keywords 1950-1969



**Figure 7.** Co-word analysis of title keywords 2000-2013



What we see is a gradual shift in granularity. There is only one cluster in 1900 focused around Milwaukee schools and something called “Bennet’s Law.” In 1950 there still is only one cluster, now minus Bennet’s Law but with a new emphasis on city schools. But in 2000 there are two distinct clusters, American education and Milwaukee Public Schools. Perhaps this reflects a cultural shift or a shifting cultural awareness in Milwaukee education.

### Conclusions

The purpose of this research was to expand our understanding of the power of domain analysis to reveal the contours of intellectual domains. Specifically, we wanted to apply the bibliometric tools of domain analysis to the base bibliography of The EMKE, an evolving online encyclopedia. This preliminary study points to the contours of the intellectual base of a culturally synergistic KOS in progress. Working with only one disciplinary category, education, is both a limitation and a methodological tactic for the present study.

We discovered typical patterns of an intellectual domain—diverse publication types, largely scattered author productivity, increased productivity at the end of the twentieth

and beginning of the twenty-first centuries, great terminological granularity but also fairly cogent thematic coherence. We also encountered some atypical results. For instance, we did not really uncover a research front, and we observed very low author productivity. We also encountered some results that cannot be explained from our study alone, such as the shift from monographic to journal venues, the shift from thesis to dissertation, and an increase in author productivity over time. Each of these latter results could be attributed as easily to the overall pattern of scholarship across time as well as to the evolution of education as a domain. Similarly, it is not clear how the domain we are viewing is different because it is derived from the base bibliography for an encyclopedia rather than from current journal or conference literature.

However, some of our results point to the encyclopedic form, which allows differentiation in the representation of a domain. This is particularly evident in the dynamic between the thematic coherence that emerged from co-word analysis, on the one hand, and in the terminological granularity also present. The domain is rich but coherent, and its core terminology has emerged from this preliminary domain analysis. The core terminology represents the extension of the domain, and the granularity represents the rich intension.

Future research using these bibliometric domain analytical techniques to evaluate the content of the other segments of the EMKE promises rich results. It also seems that additional qualitative analysis, such as Cognitive Work Analysis or ethnographic research would be promising approaches to mining the historical, sociological and geopolitical aspects of the domain of the EMKE.

## References

- De Bellis, Nicola. 2009. *Bibliometrics and citation analysis: from the Science Citation Index to cybermetrics*. Lanham, Md.: Scarecrow Press.
- Graf, Ann M., and Smiraglia, Richard. 2012. Cultural curation as classification: the evolution of the bibliography and taxonomy for The Encyclopedia of Milwaukee. In Quan-Haase, Anabel, Rubin, Victoria L., and Chaves, Debbie eds., *Information in a local and global context: proceedings of the 40th annual conference of the Canadian Association for Information Science*, Waterloo, Ontario, May 31-June 2, 2012. Available <http://www.cais-acsi.ca/>.
- Hjørland, Birger, and Albrechtsen, Hanne. 1995. Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science* 46: 400-25.
- Seligman, Amanda, and Anderson, Margo. 2012. *Encyclopedia of Milwaukee NEH grant application narrative*.
- Smiraglia, Richard P. 2012. Epistemology of domain analysis. In Smiraglia, Richard P. and Lee, Hur-Li eds., *Cultural frames of knowledge*. Würzburg: Ergon-Verlag, pp. 111-24.
- Wilson, Patrick. 1968. *Two kinds of power: an essay in bibliographical control*. Berkeley: Univ. of California Press.
- Nøjgaard, M. 1975. *Litteraturens univers: Indføring i tekstanalyse*. Odense: Odense Universitetsforlag.
- Ong, W. 1982. *Orality and literacy: The technologizing of the word*. London: Methuen.
- Turner, M. 1996. *The literary mind*. New York: Oxford University Press.
- Ulrik, A. 1965. Epic laws of folk narrative. In *The study of folklore*, ed. A. Dundes, Englewood Cliffs, N.J.: Prentice-Hall, pp. 27-45.
- Vragov, R., and I. Levine. 2007. Reviewing and revamping the double-blind peer review process. *Journal of Electronic Publishing* 10: 336-46.