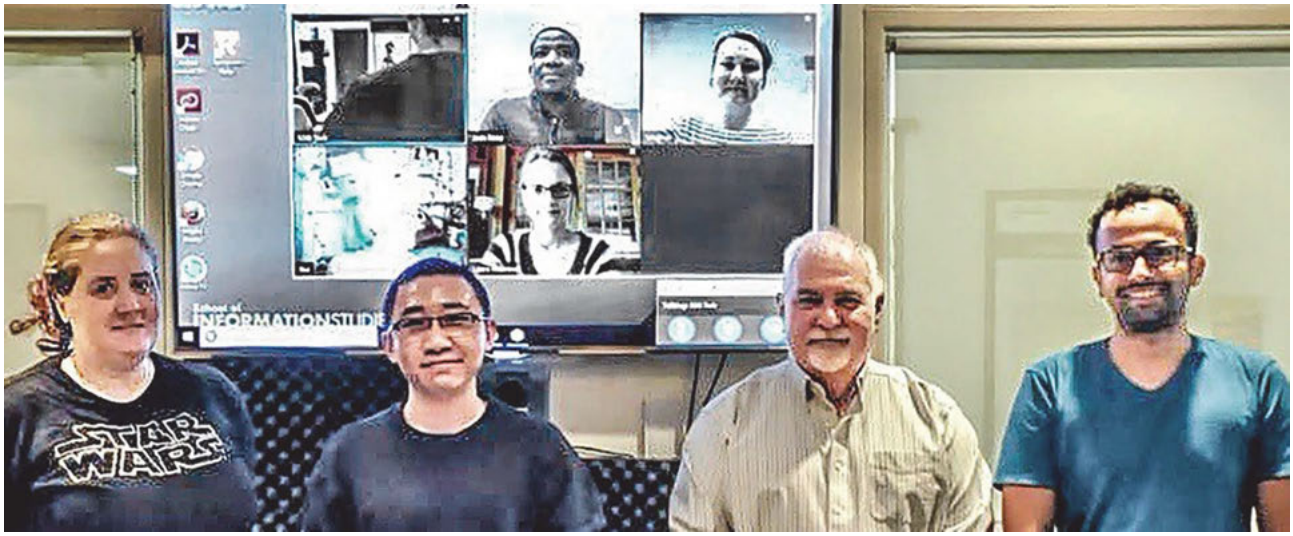
Knowledge Organization and the 2017 UDC Seminar: An Editorial

Yazeed Alhumaidan, Megan Dowell, Jean Rene, Lisa Leverett, Laura Ridenour, Vanessa Schlais, Shengang Wang and Richard P. Smiraglia

Knowledge Organization Research Group, School of Information Studies, University of Wisconsin, Milwaukee, Northwest Quad Building B, 2025 E Newport Street, Milwaukee, WI 53211, USA, smiragli@uwm.edu

Alhumaidan, Yazeed, Megan Dowell, Jean Rene, Lisa Leverett, Laura Ridenour, Vanessa Schlais, Shengang Wang and Richard P. Smiraglia. 2016. "Knowledge Organization and the 2017 UDC Seminar: An Editorial." *Knowledge Organization* 45(4): 273-280. 5 references. DOI:10.5771/0943-7444-2018-4-273.



(L-R on the screen, upper row center Jean Rene, right Meghan Dowell; lower row left Lisa Leverett, center Laura Ridenour; standing Vanessa Schlais, Shengang Wang, Richard Smiraglia, Yazeed Alhumaidan)

1.0 From Dorking 1957 to London 2017

Time like distance lends perspective to all things, thus at seventy years removed from the famous 1957 Dorking Conference that introduced faceted classification on a large scale (CRG 1997), the 2017 International UDC (Universal Decimal Classification) Seminar had facet analytical theory and all forms of the use of facets as its theme. The Conference was held in London, England on 14-15 September at Wellcome Collection. Participants gathered to celebrate "faceted analytical theory as a method for (re)constructing modern analytico-synthetic classifications and [to explore] potential fields of application for facet analysis in information organization" (http://seminar.udcc.org/2017/). The proceedings are titled *Faceted Classification Today: Theory, Technology and End Users* (Slavic and Gnoli 2017).

This editorial follows in a sequence of earlier domain analyses of international UDC seminars (Beak et al. 2014; Cai et al. 2016), the purpose which was to analyze the extension and intension of the seminar, and its overlap with the broader knowledge organization (KO) domain. The 2013 seminar was demonstrably similar to KO with regard to the constructive mix of humanist and empiricist methodologies (Beak et al. 2014, 193). But the theme of the 2015 seminar was "authority control," which seemed to bring together quite a different group of contributors, with many library-based contributors making pragmatic presentations concerning authority control and linked data, both critical for the future use of the UDC in semantic web applications (Cai et al. 2016, 402). In this editorial we present a few domain analytical visualizations to demonstrate the relative positioning of the 2017 seminar in KO. We hy-

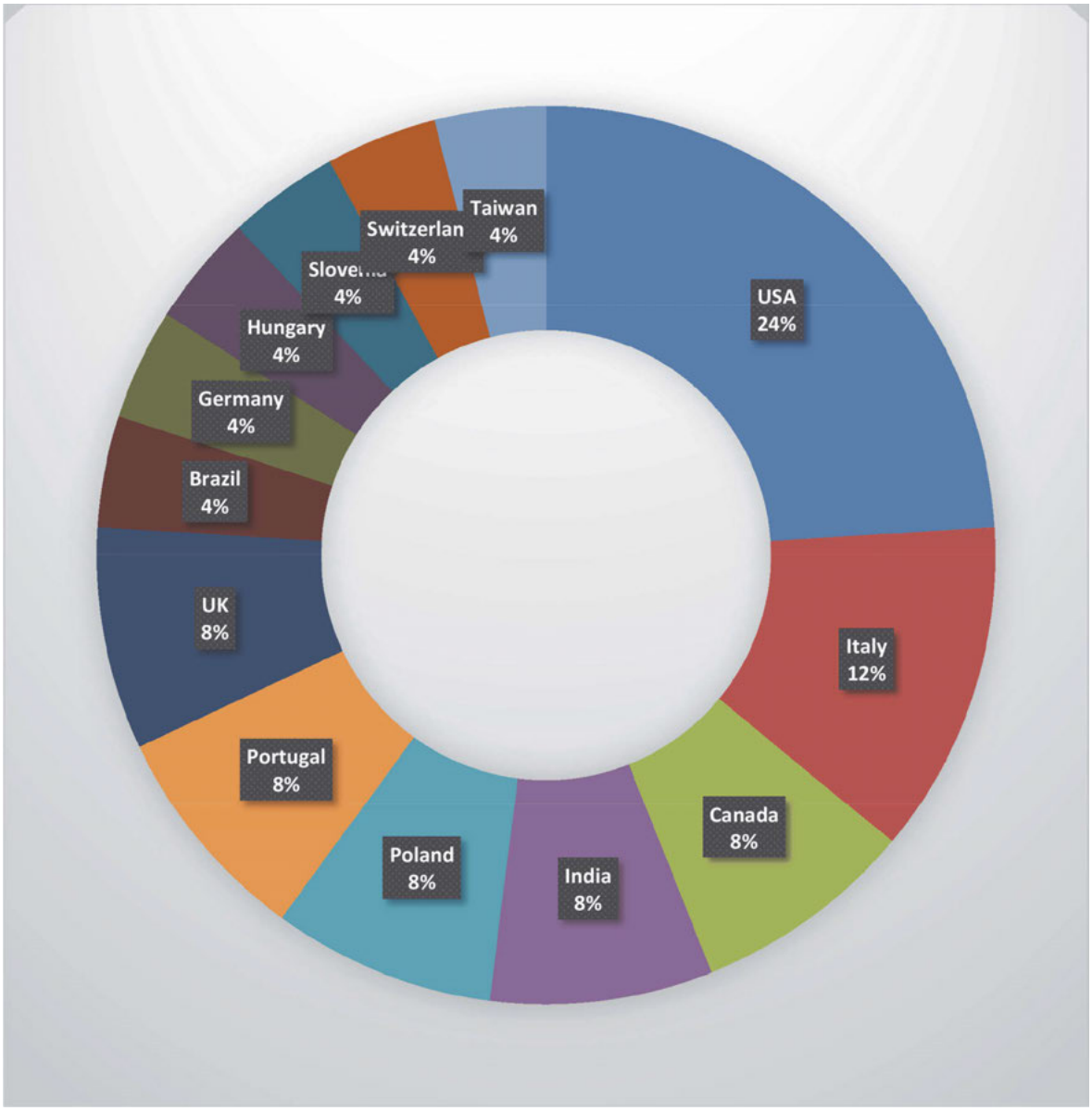


Figure 1. Countries of affiliation of contributing authors.

pothesized the focus on facet analytical theory would result in a profile very similar to that of traditional KO.

2.0 Some domain metrics

The conference consisted of seventeen formal presentations, including two keynotes by Richard Smiraglia and Vanda Broughton, and three posters. The seventeen presentations had twenty-six authors (mean 1.52; range 1-3), which aligns with prior seminars. The authors listed thirteen national affiliations as shown in Figure 1; the largest cluster was from the United States. The national affiliations are consistent with the 2013 and 2015 seminars, with the notable exceptions of the addition of Brazil and the absence of The Netherlands.

There were 377 works cited in the seventeen presentations; the mean number of citations was 26.5, with a range from 8 to 79. This represents a noticeable increase in the means from the 2013 (mean 21.4) and 2015 (mean 10.07) seminars, which is consistent with this seminar's greater emphasis on historical documents.

The year of works cited ranged from 1884 to 2017 and the mean age of cited work was 37.8 years again reflecting the historical bent of this seminar. The distribution of dates of publication of works cited is visualized in Figure 2.

Most works cited were published in 2000 or later, which is comparable with the earlier seminars with regard to relative contemporaneity. Mean age of cited work by contributing author is visualized in Figure 3.

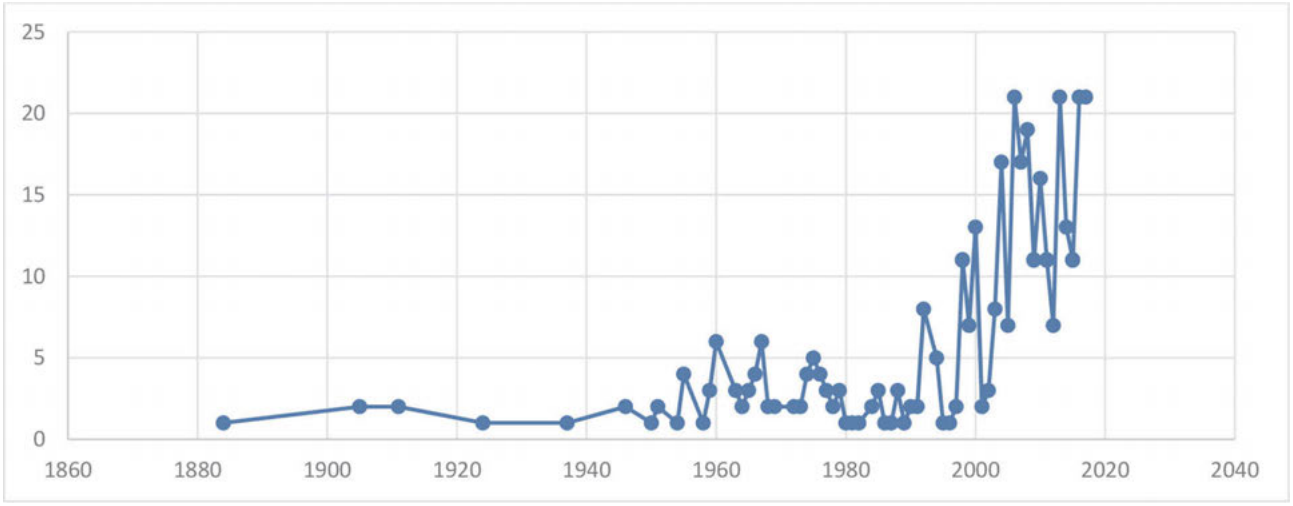


Figure 2. Year of works cited.

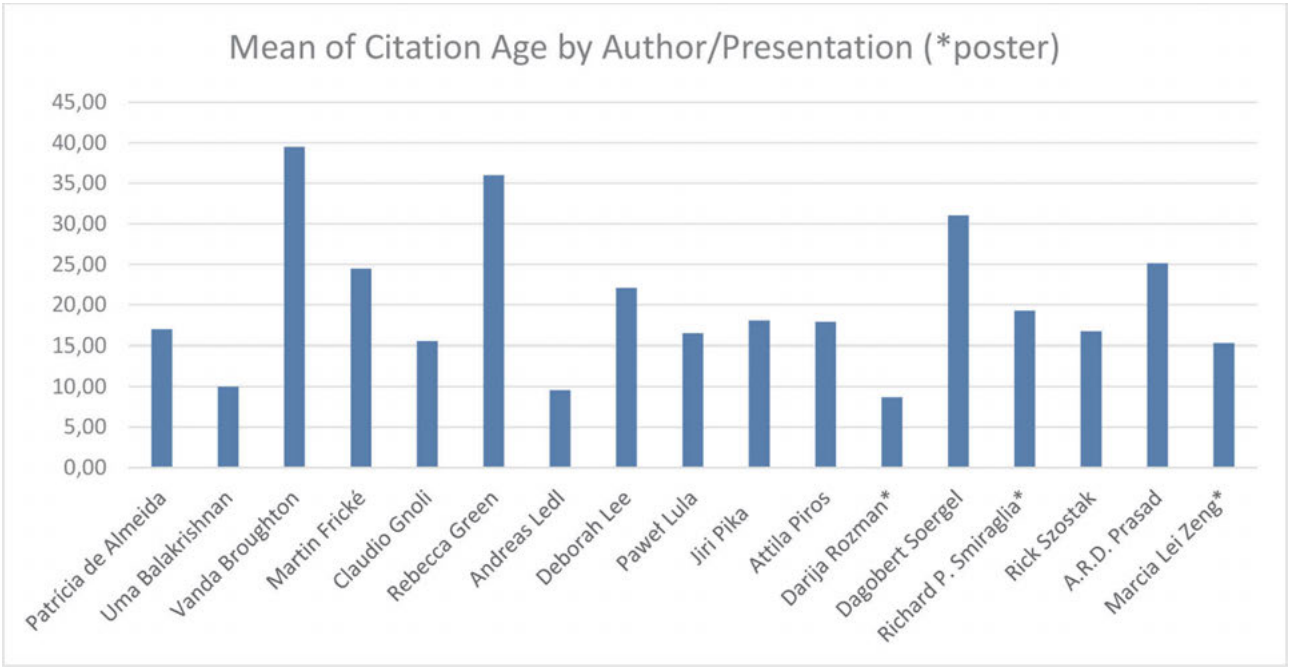


Figure 3. Age of work cited by contributor.

This visualization demonstrates quite clearly the recency of most cited works, which occur in empirical research papers. Thus we can conclude the addition of historical references, especially in the keynote presentations, lead to the extension of the range of dates of works cited. Otherwise, the date profile of works cited is similar to earlier seminars and to KO in general. Twenty-six works were cited two or more times; eleven were cited three or more times (Table 1).

There are no surprises on this list; the most cited works are those that are considered core writings in the use of facets for knowledge organization (Smiraglia 2017).

One hundred thirty-two of the 377 citations were to journal articles (35%), 92 were to monographs (24%), 41 (11%) were to papers in conference proceedings. The remaining 66 citations (30%) were to mid-twentieth century technical reports or current online resources, such as blog posts. The distribution of main sources is typical in knowledge organization and prior seminars, but the growing proportion of citations to web resources is notable. Conferences that predominate are ISKO International Conferences, NASKO (North American Symposium on Knowledge Organization, the biennial conference of ISKO-Canada/United States), and International UDC

Ranganathan, S. R. 1967. *Prolegomena to library classification.* 3rd ed. London: Asia Publishing House.	5
Hjørland, Birger. 2013. Facet analysis: the logical approach to knowledge organization. *Information Processing and Management* 49: 545-57.	4
La Barre, Kathryn. 2010. Facet analysis. *Annual Review of Information Science and Technology* 44: 243-84.	4
Vickery, Brian. 1960. *Faceted classification: a guide to construction and use of special schemes.* London: Aslib.	4
Aitchison, Jean, Alan Gilchrist and David Bawden. 2000. *Thesaurus construction and use.* 4th ed. London: ASLIB.	3
Broughton, Vanda. 2004. *Essential classification.* London*:* Facet.	3
Classification Research Group. 1955. The need for a faceted classification as the basis of all methods for information retrieval. *Library Association Record* 57, no. 7: 262-68.	3
Gnoli, Claudio. 2011. Facets in UDC: a review of [the] current situation. *Extensions and Corrections to the UDC* 33: 19-36.	3
Spiteri, Louise. 1998. A simplified model for facet analysis: Ranganathan 101. *Canadian Journal for Information and Library Science* 23: 1-30.	3
Vickery, Brian. 1966. *Faceted classification schemes.* Rutgers Series on Systems for the Intellectual Organization of Information 5. New Brunswick, NJ: Graduate School of Library Science at Rutgers University.	3
Vickery, Brian. 1975. *Classification and indexing in science.* 3rd ed. London: Butterworths.	3

Table 1. Most-cited source works.

seminars. The most productive journals are shown in Figure 4.

Cataloging & Classification Quarterly, *Knowledge Organization*, *Journal of Documentation*, and the *Journal of the Association for Information Science and Technology* (and its predecessors) and the *Annual Review of Information Science and Technology* are typically found in KO research publications. As in prior seminars, the extension of the domain of the conference is clearly illustrated by the top-ranked *Extensions & Corrections to the UDC*.

3.0 Co-word analysis

Co-word analysis was conducted using titles and abstracts of articles of the UDC proceedings. Titles and abstracts were pre-processed to remove stopwords and meta-terms about the paper and study (e.g. "this paper"), and stemmed using the built-in algorithm in Sci2. The stemmed terms were processed for co-occurrence using Sci2's build in word co-occurrence algorithm. The resulting file was loaded into Gephi and reduced to include twenty nodes, or 3.29% of the total number of stemmed terms. As the focus of this analysis was directed at content-bearing terms, bigrams and trigrams represented in tables 2 and 3 include those most frequently occurring, as detected using textalyser.net.

The visualization in Figure 5 represents the most interconnected head of the long-tail of the corpus, which is the majority of single stemmed terms that co-occur in more than one document. The larger nodes (e.g., "use," "facet," "classif") show terms that co-occur with highest frequency. More important are the lines, or edges, connecting the nodes—these represent the weight of the frequency of co-occurrence of pairs of terms (such as "knowledge" and "organization," or "UDC" and "seminar").

4.0 Author co-citation analysis

The authors who were most cited by seminar participants are named in Table 4.

These authors' names were used to create a co-citation matrix, and occurrences of author co-citation among the seminar papers were documented manually. The matrix was entered into IBM-SPSS™ and then used to generate the multi-dimensionally-scaled (MDS) plot show in Figure 6. Data cluster in two large groups, roughly in the upper and lower hemispheres of the plot. The upper cluster represents classical writing about facet analytical theory and faceted classification. The lower cluster represents the research front of the seminar, which interestingly is anchored by Broughton, whose work on the influence of facet analytical theory on information retrieval spans several decades and is perhaps representative of the catalytical turning point driving the research front.

The same matrix was used to create a network visualization using Vosviewer 1.6.5, shown in Figure 7. Work by Gnoli, La Barre, Slavic, Mills, Vickery and Ranganathan is cited by participants as representing the dense core of the conference, faceted classification. The edges show the strength of the associations, particularly with regard to

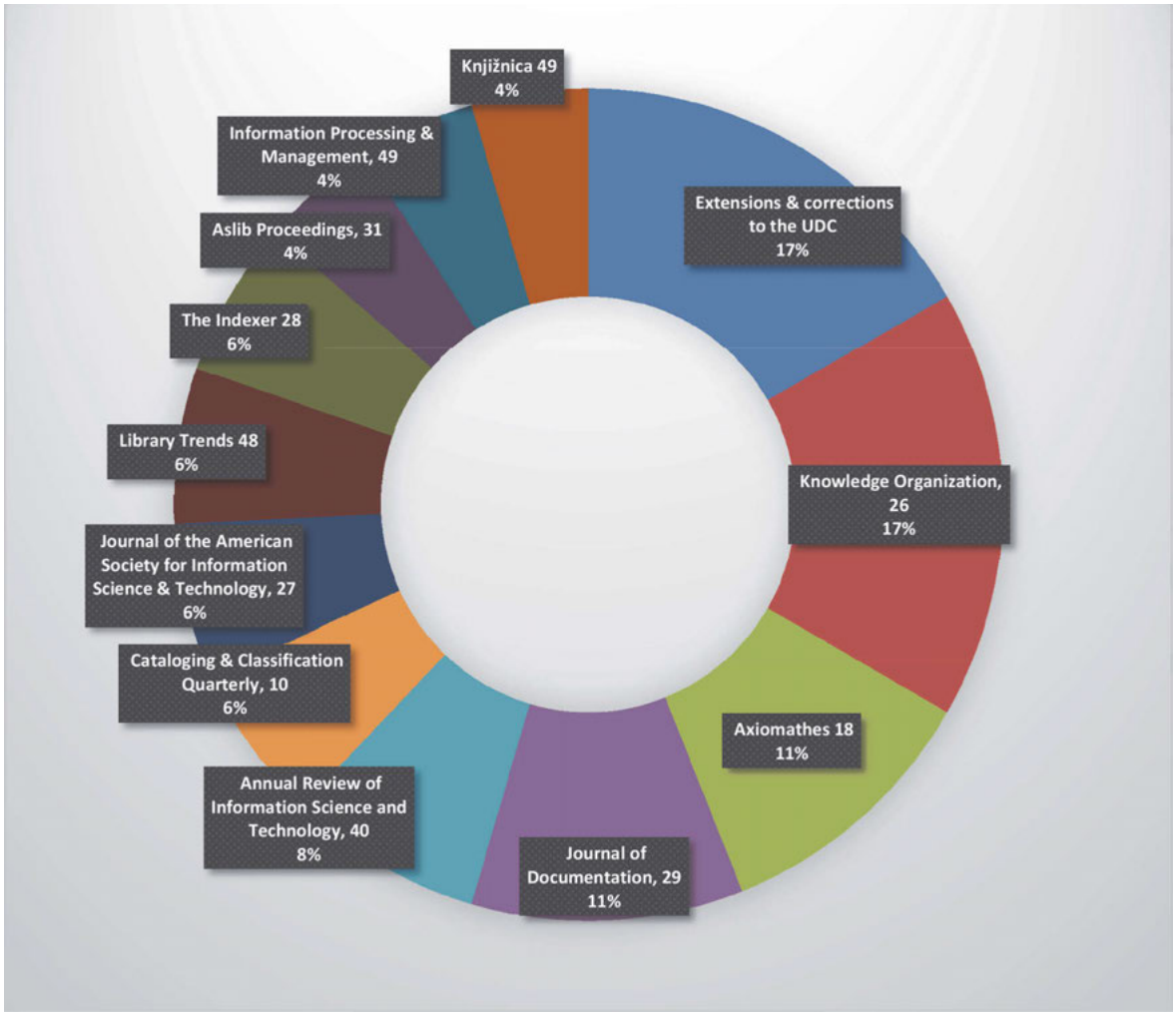


Figure 4. Journals cited most frequently.

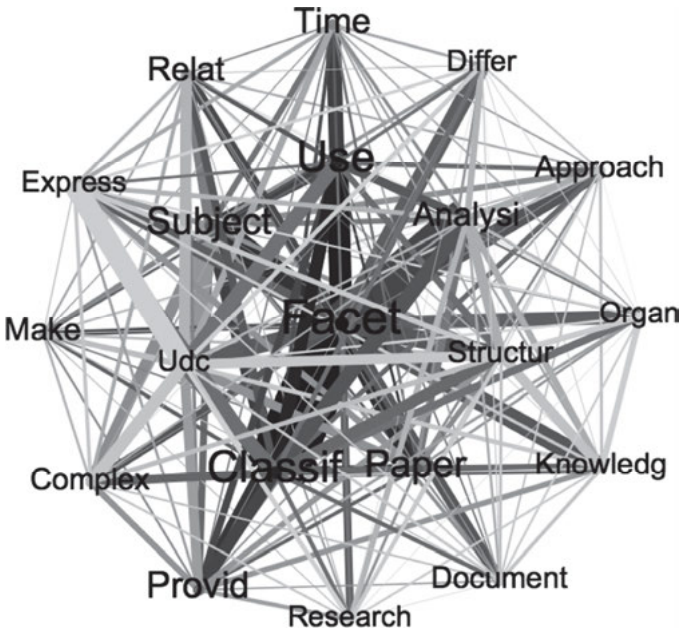


Figure 5. Visualization of relationships (co-word analysis of stemmed terms).

Expression	Expression count	Frequency	Prominence
facet analysis	16	0.50%	34.8
faceted classification	9	0.30%	46.8
sub facets	6	0.20%	64.7
entity relationship	5	0.20%	25
knowledge organization	5	0.20%	56.7
relationship modelling	4	0.10%	23.8
analytico synthetic	4	0.10%	31.2
udc expressions	4	0.10%	58.9
semantic frame	4	0.10%	72.9
knowledge representation	3	0.10%	24.7
udc classmarks	3	0.10%	46.7
complex subjects	3	0.10%	48.8
pre coordinated	3	0.10%	51.4
complex udc	3	0.10%	51.5
subject index	3	0.10%	53.2
faceted systems	3	0.10%	54.3
organization systems	3	0.10%	60
music classification	3	0.10%	65.9
determination of	3	0.10%	66.8
classification which	3	0.10%	66.9
frame analysis	3	0.10%	73.5
subject classification	3	0.10%	85.6

Table 2. Content-bearing bigrams occurring three or more times.

Expression	Expression count	Frequency	Prominence
entity relationship modelling	4	0.10%	23.8
knowledge organization systems	3	0.10%	60.1
semantic frame analysis	3	0.10%	73.6
basic faceted classification	3	0.10%	87.4
entity centric approach	2	0.10%	9.4
basic concepts classification	2	0.10%	13.1
principle of compositionality	2	0.10%	15.3
systems faceted classification	2	0.10%	18
phenomenon based classification	2	0.10%	36.6
model for classifying	2	0.10%	65.5
mode of search	2	0.10%	86.6
foci across facets	2	0.10%	88.7
use and exchange	2	0.10%	92.8

Table 3. Content-bearing trigrams occurring twice or more.

Broughton, Gnoli and Vickery, whose work collectively is closely associated as core by citing participants. The density of associated relationships on the left side of the diagram relate to the use of facet analytical theory in classification for information retrieval.

5.0 Faceted classification for information retrieval

Over time the importance of the international UDC seminar to the domain of knowledge organization has been clear (Smiraglia 2017). Although participation has varied over time, especially with regard to seminar themes, this latest iteration is clearly in line with prior seminars as well as with the domain of KO in general. The rich link between this seminar and the history of the facet analytical theory movement serves both to root the research presented here in the core of KO and to provide a backdrop for catalyzing new developments in the use of facets for information retrieval. This is clear from every part of our analysis from the age of works cited to the list of most-cited authors to the thematic clusters revealed in co-word and author co-citation analysis, thus neatly tying together the various observations.

Author	Citation frequency
Broughton, V.	14
Gnoli, C.	14
Szostak, R.	12
Vickery, B.	12
Giunchiglia, F.	10
Ranaganthan, S. R.	9
Hjørland, B.	8
La Barre, K.	7
Frické, M.	6
Slavic, A.	6
Smiraglia, R.	5
Soergel, D.	5
Foskett, D. J.	4
Gardin, J.-C.	4
McIlwaine, I. C.	4
Mills, J.	4

Table 4. Authors most frequently cited four times or more by participants.

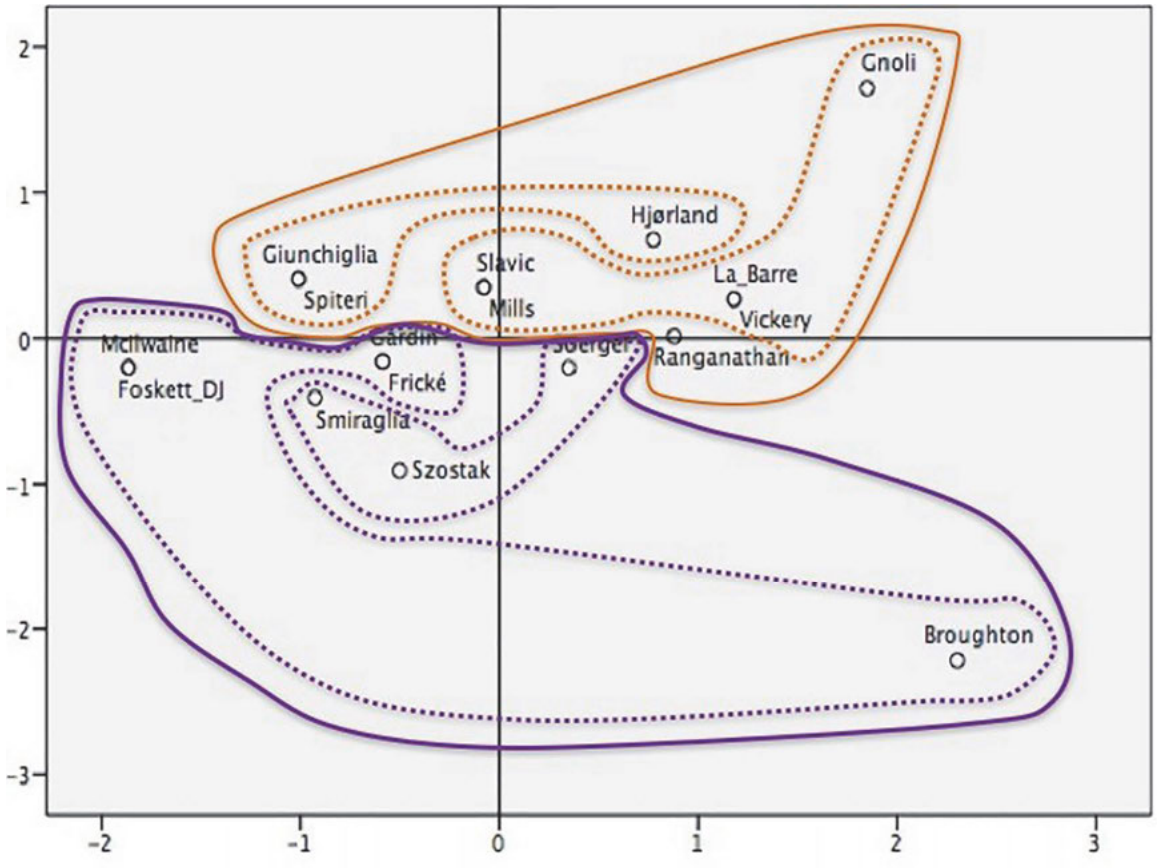


Figure 6. MDS plot (stress = .08 R2 = .97) of internal author co-citation.

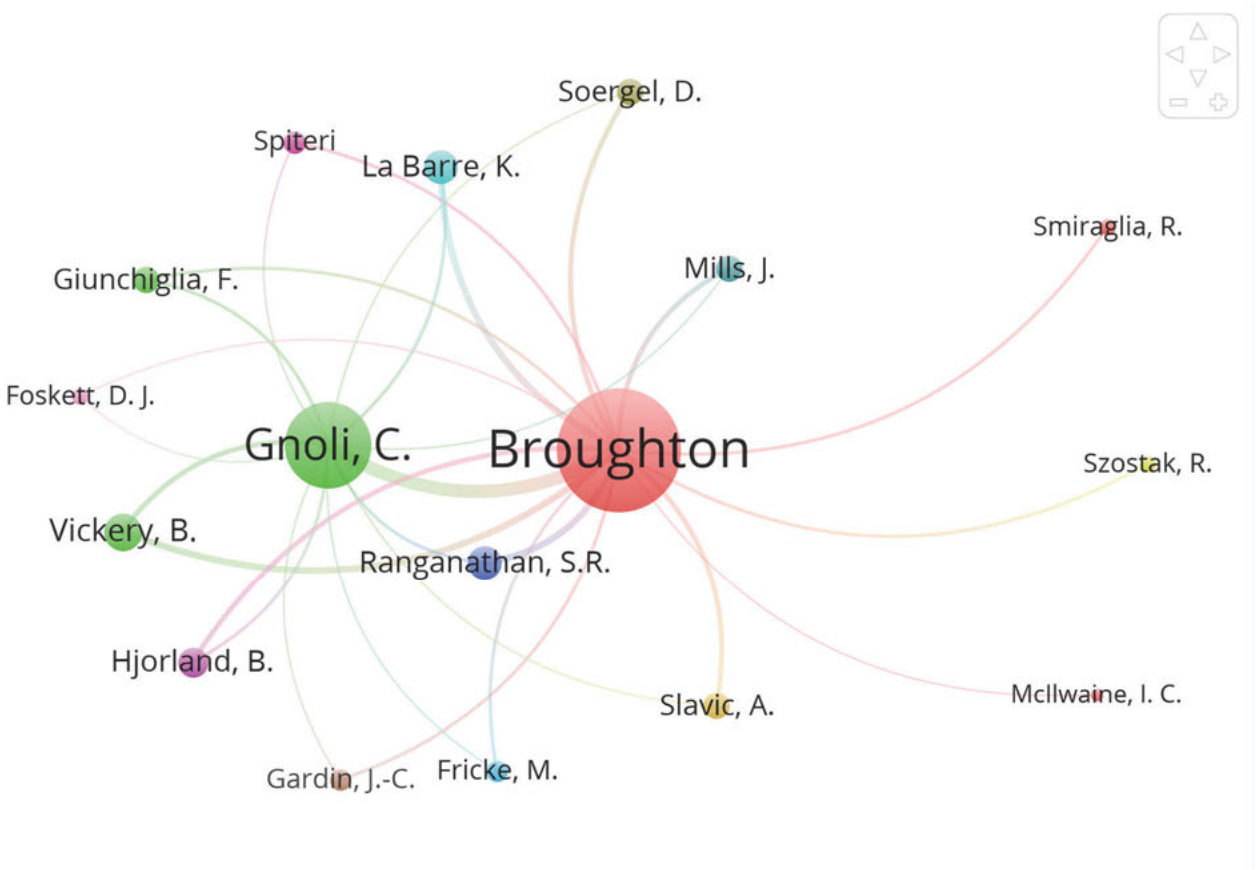


Figure 7. Vosviewer 1.6.5 network diagram of internal author co-citation.

References

Beak, Jihee, Inkyung Choi, Sukwon Lee, Hyoungjoo Park, Laura Ridenour and Richard P. Smiraglia. 2014. "Knowledge Organization and the 2013 UDC Seminar: An Editorial." *Knowledge Organization* 41: 191-4.

Cai, Xin, Melissa Davey Castillo, Musa Dauda Hassan, Ann Graf, Hyoungjoo Park and Richard P. Smiraglia. 2016. "Knowledge Organization and the 2015 UDC Seminar: An Editorial." *Knowledge Organization* 43: 395-402.

Classification Research Group (CRG). (1997 [1955]) The need for a faceted classification as the basis of all methods of information retrieval: Memorandum of the Classification Research Group: received, for information, by the L. A. Library Research Committee, May 1955. In: *From classification to "knowledge organization: Dorking revisited or "past is prelude"*. Edited by A. Gilchrist. The Hague: FID, 1997. (FID 714), 1-9.

Slavic, Aida and Claudio Gnoli, eds. 2017. *Faceted Classification Today: Theory, Technology and End Users: Proceedings of the International UDC Seminar, 14-15 September, London, United Kingdom*, organized by UDC Consortium, The Hague. Würzburg: Ergon Verlag.

Smiraglia, Richard P. 2017. "Facets as Discourse: How Facets and Facet Analytical Theory Reveal Cultural Dimensions in 21st Century Knowledge Organization Systems." In *Faceted Classification Today—Theory, Technology and End Users: Proceedings of the International UDC Seminar, 14-15 September 2017, London, United Kingdom*, ed. Aida Slavic and Claudio Gnoli. Würzburg: Ergon Verlag, 3-22.