Classifying KOSs: The Use of Dahlberg's ICC in the DANS KOS Observatory (KOSo)

Coen, Gerard; Smiraglia, Richard P.; Doorn, P.K.; Scharnhorst, A.M.

2019

Classifying KOSs: The Use of Dahlberg's ICC in the DANS KOS Observatory (KOSo)

(ISKO Low Countries)

Gerard Coen, Richard P. Smiraglia, Peter Doorn, Andrea Scharnhorst

Abstract

This paper describes the theoretical framework and methodological workflow to classify Knowledge Organisation Systems (KOSs). As part of this research a KOS Observatory has been set up which currently contains about 132 KOS from various disciplines, primarily social sciences, humanities and life sciences. The DANS KOS Observatory (KOSo) partly integrates and partly complements existing KOS registries. This research project is innovative on two different grounds. It represents a first attempt to consistently classify KOS using the Information Coding Classification. Second, the observatory is set up in a way to support long term data archiving (including the archiving of related KOS), hence it supports data curation both during research and during long term preservation. With the KOS Observatory (KOSo) reconstruction and study of the evolution of KO (Knowledge Organization) systems becomes possible, but it is also of immediate practical importance for researchers searching for relevant KOS for their research question and for data managers at research data archives in their routines to check the quality of submitted data collections. This paper focus on the workflow to classify KOSs in the KOSo.

1.0 Classifying KOSs

The purpose of this research is to explore the process of classifying the classifiers—that is, to explore the process of classifying knowledge organization systems (KOSs). Unlike the classification of documents based on judgments of aboutness, classifying KOSs requires sensitivity both to domain-specificity and to form of KOS. KOSs are widely variable, ranging from functional systems for organizing information, to conceptual systems describing domains, reflecting basic human mental patterns, and increasingly fueling information architectures in the semantic web (Mazzocchi 2018). Thus the classification of KOSs, which may provide conceptual, thematic and semantic orderings of any potential domain, itself must reflect the classification of knowledge at a meta-level. To that end, we report our use of the Information Coding Classification (ICC) designed by Ingetraut Dahlberg (1982a and b) for just that purpose. Specifically, we have applied the ICC to the growing corpus of the Knowledge Organization System Observatory (KOSo) being curated by the DANS division of the Royal Netherlands Academy of the Arts and Sciences. In this paper we report our experience "classifying the classifiers" using specific cases from the KOSo.

2.0 DANS, FAIR and the development of KOSo

Since 2016, European research institutions, particularly those tasked with data stewardship, have been making a concerted effort to align to the FAIR data principles (https://www.force11.org/group/fairgroup/fairprinciples), which aim "to facilitate knowledge discovery by assisting humans and machines in their discovery of, access to, integration and analysis of, task-appropriate scientific data and their associated algorithms and workflows." This is achieved by making data FAIR - Findable, Accessible, Interoperable, Reusable (Wilkinson, et al., 2016).

DANS (Data Archiving and Networked Services) is the Dutch institute responsible for the stewardship of national research information. It is affiliated to both the Royal Netherlands Academy of the Arts and Sciences (Koninklijke Nederlandse Akademie van Wetenschappen or KNAW) and the Netherlands Organisation for Scientific Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek or NWO). In an effort to lean into its role as an advocate for FAIR, as well as in association with the KNAW Visiting Professor Richard Smiraglia's formal programme, in 2017 DANS began the Knowledge Organization Systems Observatory (KOSo). The KOSo sets out to explore the interoperability aspect of the FAIR data principles while simultaneously developing more insight into the history, evolution and mutual dependencies of the different domain-specific KOSs with which DANS sees itself confronted. The KOSo, now in its second year, contains over 132 KOSs from the Social Sciences and Humanities (SSH) and the Life Sciences (LS), which have been been described using a broad variety of attributes. Furthermore, the KOSo has been annotated using both Dahlberg's International Coding Classification (ICC) and the NARCIS Classification, a proprietary research classification schema used by DANS for the Dutch research information portal of the same name, NARCIS.

3.0 Methodology

3.1 Comparative classification

Building on the work of Szostak and Smiraglia (2017a; 2017b; 2018) and Coen and Smiraglia (2019) this paper applies methods of comparative classification to the KOSo. Comparative classification is a recent research stream in the field of Knowledge Organization (KO) which uses two or more classifications experimentally to designate content.

Annotating the KOSo using two classifications provides both practical benefits and theoretical insights. Practically, the use of the NARCIS Classification creates a first step towards

integrating the KOSo within existing DANS services, namely NARCIS. In this way DANS actively works towards the FAIRification of these resources, and the advancement of its own portfolio of services. On an academic level, applying the technique of comparative classification to the observatory allows us to generate empirical evidence about the coverage of conceptual content, including precision, the population of the classes, and also, the economy of classification provided by both KOSs. In doing so, the cultural dependence of these classifications becomes apparent (Beghtol, 2010). That is to say that we can gain insight into the impact of varying forms of warrant on the development of these KOSs, most of which were designed for the ordering of research data. Furthermore, the suitability of either NARCIS or the ICC as a classification schema for the KOSo can be assessed in this way.

Specifically, we gather data into the KOSo about extant KOSs according to this schema:

Identifier
Schema Name/Title
Creator(s)/Curator(s)
Maintenance organization
Place: Publisher, date.
Summary / Abstract
Format(s)
Language
Physical Location
Online Location
Earlier versions (editions) …
History of versioning:
Version Notes:
Last Updated:
Number of terms in system
Phenomena included
Disciplines included
Direct domains included
Related to:
BARTOC Link
ICC Classification
NARCIS Classification
KOS Types Vocabulary

We also are concerned with the problems of instantiation and scheme-versioning. The objective is to create, for the KO community, the ability to track and observe KOSs as they mutate over time and across implementations. Tennis (2002; 2007; 2012; 2015; 2016) has

been the most prominent catalyst for the KO community, joined by Fox (2016), Lauruhn and Gorth (2016), Lee (2016), Salah et al. (2012), Scharnhorst et al. (2012; 2016), Scharnhorst and Smiraglia (2012) and Smiraglia et al. (2013). Expressions from the KO and semantic web (SW) communities were brought together in two KNOWeSCAPE workshops held in Amsterdam (2015) and Malta (2017) to prioritize objectives for an observatory for KOSs.[1] In the KOSo implementation then, we have taken care to track all extant instantiations of each KOS, thus creating via our metadata schema above an instantiation matrix using system identifiers. Thus the matrix using these identifiers visualizes the evolution of a specific KOS, see the example of the *Art & Architecture Thesaurus* (*AAT*) (Figure 1)

Identifier	Schema Name/Title	Creator(s)/Curator(s)	Maintenance organization	Place: Publisher, date.
KOS.SSH.2	Art & Architecture Thesaurus; AAT	Toni Petersen	Getty Art History Information Program	New York : Oxford University Press, 1990
KOS.SSH.2.1	Art & Architecture Thesaurus; AAT	Toni Petersen	Getty Art History Information Program	New York : Oxford University Press, 1992
KOS.SSH.2.2	Art & Architecture Thesaurus; AAT	Toni Petersen	Getty Art History Information Program	New York : Oxford University Press, 1994
KOS.SSH.2.3	Art & Architecture Thesaurus; AAT	Toni Petersen	Getty Art History Information Program	New York : Oxford University Press, 1994
KOS.SSH.2.4	Art & Architecture Thesaurus; AAT	Toni Petersen	Getty Art History Information Program	New York : Oxford University Press, 1994
KOS.SSH.2.5	Art & Architecture Thesaurus; AAT	Currently, Patricia Harpring Editors are Jonathan Ward The technical team is led b	Getty Art History Information Program	The J. Paul Getty Trust
KOS.SSH.2.6	Nederlandstalige AAT	Reem Weda	RKD-Nederlands Instituut voor Kunstgesch	The Hague : RKD- Nederlands Instituut voor Kur
KOS.SSH.2.7	AAT-Deutsch	Frau Prof. Hagedorn-Saupe	Institut für Museumsforschung (& Getty Art History Information Program)	Berlin : Institut für Museumsforschung, 2014
KOS.SSH.2.8	TAA; El Tesauro de Arte & Arquitectura	n/a	El Centro de Documentación de Bienes Pa (& Getty Art History Information Program)	Santiago : El Centro de Documentación de Bien

***Figure 1*. KOSo matrix for AAT.**

The schema and metadata matrix visualization is discussed in greater detail with illustrations in Szostak *et al*. (2018) and Coen and Smiraglia (2019).

3.2 Classifying KOSo

Of course, it was essential to use knowledge organization tools to organize our data; early in the project we understood the potential importance of both metadata descriptors for form and type of KOS--e.g., thesaurus, taxonomy, etc.--as well as classification to place the entire observatory in intellectual space--e.g, thesaurus of art descriptors, taxonomy of agriculture, etc. Thus, we decided to implement the ICC, described by Dahlberg (most clearly in 2008) as a classification of knowledge fields, instead of disciplines, organized according to the theory of integrative levels (163): “the law, that one level presupposes the next one and that they integrate each other.” Thus, fields relating to matter and energy precede those relating to aggregated matter, which follows and thus is integrated, in the same manner that fields relating to human beings precede those relating to societal beings, etc. Dahlberg also describes the ICC as “fully-faceted” because its matrix-like structure is based on the intersection of the nine integrative levels with a set of nine general form categories (e.g., theories, application of methods, etc.), which constitute a technique she identifies as the “Systematifier,” thus faceting is accomplished by systematic placement of knowledge fields in the matrix requiring no synthesis of notation. Finally, the ICC allows synthesis or “combinatory functions” (172) representing the expression of phase relationships. Thus, to give some examples, “history of mathematics would receive the notation 121:698 or sociology of education 561:618” and “if 824 archvistics, [and] 82.32 information analysis, then 824.27 archival manuscripts 824.27.32 [and] information analysis of archival manuscripts.” For KOSs ICC is used in a synthetic format together with Dahlberg’s (1999) KO Literature Classification (KOLC), thus rendering for each KOS a domain-centric term faceted with a KOS-form term. For KOSo purposes the following forms are used in combinatorial format (192-3):

03 Glossaries, Vocabularies, Terminologies in Knowledge Organization
031 General Glossaries in Knowledge Organization
032 Glossaries Containing Knowledge Organization Sections
033 Free
034 Terminology of Universal Systems
035 Free
036 Terms and Glossaries in Special Knowledge Organization Fields

037 Terminology of Special Classification and Indexing Systems
038 Terms and Glossaries in Knowledge Organization Application Fields
039 Free
04 Universal Classification Systems
041 Library Classifications Systems
…
048 Other Universal Classification Systems and Thesauri
048-1 Classification Systems and Thesauri in Logic, Mathematics and other Formal Sciences
048-2 Classification Systems and Thesauri in Physics, Chemistry, Electronics, Energy
048-3 Classification Systems and Thesauri in Astronomy, Geosciences, Geography, Mining
048-4 Classification Systems and Thesauri in Biological, Veterinary Science, Agriculture, Food Sciences, Ecology
048-5 Classification Systems and Thesauri in Human Biology, Medicine, Psychology, Education, Labour, Sports, Household
048-6 Classification Systems and Thesauri in the Sociology, Politics, Social Policy, Law, Area Planning, Military Science, History
048-7 Classification Systems and Thesauri in Economy, Management Science, Mechanical Engineering, Building, Transport
048-8 Classification Systems and Thesauri in Science of Science, Information Science, Computer Science, Communication Science, Semiotics
048-9 Classification Systems and Thesauri in Language, Literature, Music, Arts, Philosophy, Religion

It can be seen, then, that the divisions of class 048 “classification systems and thesauri” follow the integrative levels of the ICC matrix. Thus, in KOSo, the ICC classification for the *AAT* is 048-944; 048-945, where 944 represents Malerei und Graphik (painting and graphics) and 945 Baukunst (architecture). Apart from this usage, we have not implemented the combinatory functions as of the writing of this paper.

Additionally, as we contemplated positioning the KOSo data among DANS online resources we began to classify domains using the DANS NARCIS Classification (described more fully in Coen and Smiraglia 2019). The NARCIS Classification (https://www.narcis.nl/classification/Language/en) is designed as a framework to represent the research foci of the Dutch national research infrastructure. In the cases that follow we demonstrate the efficacy of ICC used with KOLC as compared to the NARCIS Classification.

4.0 Results

This work of ‘classifying the classifiers’ (as we call it) is still a work in progress. This paper provides a view on the assignment of classifications for five specific cases (two KOSs from the Life Sciences, and three from SSH domains) using the NARCIS Classification and the International Coding Classification. The classification was performed using the standard symbols in both classification systems. While we are aware that the ICC provides the possibility to synthesize symbols using “combinatory functions” (Dahlberg 2008, 171), whereby for example 12 “Mathematics”, 05 “Human Area” and 08 “Science and Information” can become 12.05.8 ‘Scientists in Mathematics’, such synthesis has not been elaborated in this study.

Overall, at the time of writing, the Life Sciences KOSo contains twenty-four KOSs which have a combined seventy-four instantiations. Of those seventy-four instantiations, there are seven instantiations which have been assigned two ICC symbols. When assigning NARCIS symbols to the same collection, there are eight instantiations which are not easily classified with one symbol. Of these eight, one instantiation has been assigned two NARCIS symbols, and six instantiations have been assigned three NARCIS symbols. The remaining instantiation is not classified.

The Social Sciences and Humanities KOSo contains 105 KOSs which have combined 298 instantiations. Of these 298 instantiations, thirty-eight have two or more ICC symbols, fourteen have two NARCIS symbols. A further twenty-three have no ICC symbol, and twelve have no NARCIS symbol. These simple metrics help us to get a better appreciation of the expressivity and precision of the classifications, as demonstrated in the following examples. Further work to understand the benefit of using synthesis for the representation of the KOSs in the KOSo is still in its early stages given its complex nature.

4.1 Case results from comparison of ICC and NARCIS

Figures 2 through 7 below illustrate the comparative classifications we were able to assign to five cases from KOSo.

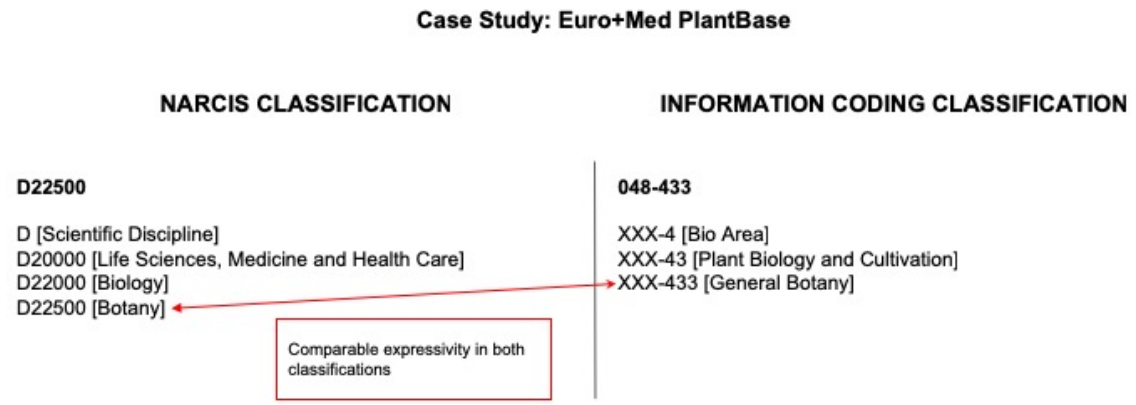

Figure 2. **Classification of Euro+Med PlantBase.**

Figure 2 shows the Euro+Med PlantBase. It provides an online database and information system for the vascular plants of Europe and the Mediterranean region (http://www.emplantbase.org/home.html). In this case both classifications offer a precise location; concepts relating to well established scientific disciplines, particularly from the Natural Sciences, are explicitly expressed in both systems.

Case Study: Centre for Agriculture and Bioscience (CAB) Thesaurus

NARCIS CLASSIFICATION	INFORMATION CODING CLASSIFICATION
D10000 D [Scientific Discipline] D10000 [Science and technology]	**048-4** XXX-4 [Bio Area]
D20000 D [Scientific Discipline] D20000 [Life Sciences, Medicine and Health Care]	**048-5** XXX-5 [Human Area]
D40000 D [Scientific Discipline] D40000 [Law and Public Administration]	**048-6** XXX-6 [Socio Area]
D60000 D [Scientific Discipline] D60000 [Social Sciences]	It is unclear how to classify a KOS with such a broad coverage.

Figure 3. **Classification of the Centre for Agriculture & Bioscience (CAB) Thesaurus.**

Figure 3 shows the Centre for Agriculture and Bioscience (CAB) Thesaurus. It offers coverage of Pure and Applied Life Sciences, Technology and Social Sciences. It has a total number of terms approaching 2.8 million.

Coverage of the CAB Thesaurus

1. GENERAL
- common terms - general processes, properties and characteristics
- named regions and countries
- research and methodology
- mathematics, statistics and computer science
- communication and information

2. PHYSICAL SCIENCES
- physics
- Chemistry

3. EARTH SCIENCES
- geology
- geomorphology
- soil science
- hydrology
- meteorology and climatology

4. LIFE SCIENCES
- biology
  - microbiology
  - botany
  - zoology
  - cytology
  - embryology
  - molecular biology
  - biochemistry
  - physiology
  - genetics
  - biological structure and form
  - taxonomy
- ecology
- behaviour and psychology
- organisms

5. APPLIED SCIENCE AND TECHNOLOGY
- health and pathology
  - diseases
  - pathogenesis and epidemiology
  - health and health protection
  - pharmacology and toxicology
- applied human and animal nutrition
- applied genetics and breeding
- agriculture, forestry and fishery
- environment and natural resource management
- technology and engineering
  - materials and equipment
  - materials handling and processing
  - transport
  - energy and power
- food science

6. SOCIAL SCIENCES AND HUMANITIES
- social sciences
  - education
  - sociology
  - social welfare
  - policy, law and legislation
  - economics
- culture and humanities
  - leisure, recreation and tourism

Source: https://www.cabi.org/cabthesaurus/mlwtk.exe?vi=coverage

Figure 4. **Coverage of the CAB Thesaurus.**

In Figure 4 an overview of the higher level disciplines and domains covered by the CAB Thesaurus can be seen. Its broad coverage makes it a challenge to classify using either ICC or the NARCIS Classification. For either classification a minimum of three symbols is needed.

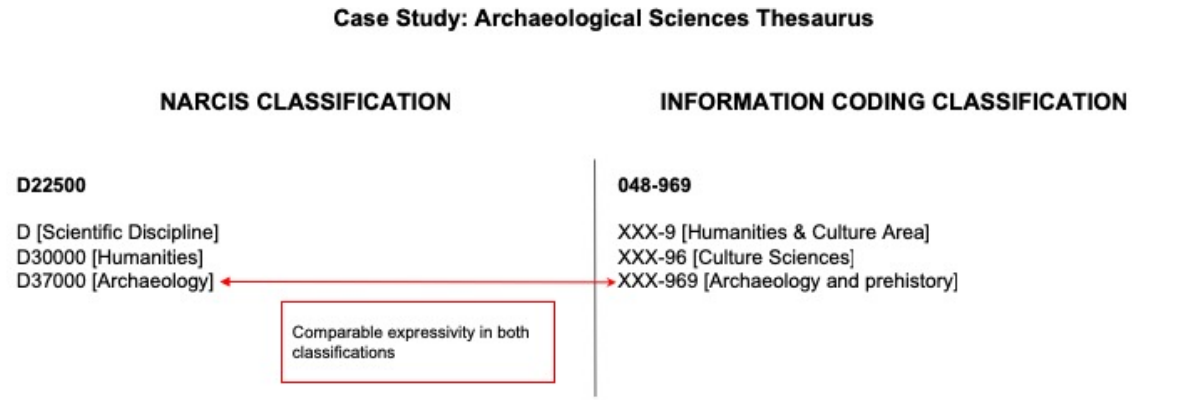

Figure 5. **Classification of Archaeological Sciences Thesaurus.**

The Archaeological Sciences Thesaurus provides terminology used for recording the techniques, recovery methods and materials associated with archaeological sciences (e.g. tree-ring analysis, modification state, pathology). Again for this KOS both classifications offer comparable expressivity.

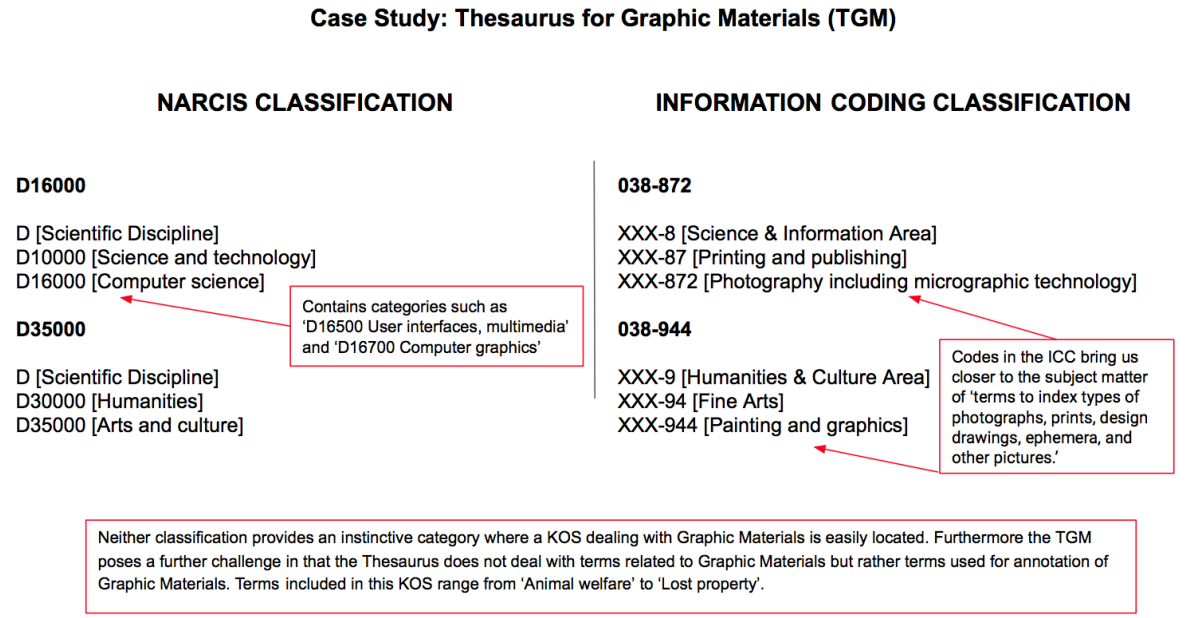

Figure 6. **Classification of Thesaurus for Graphic Materials (TGM).**

The Thesaurus of Geographic Names (TGN) is one of the Getty Vocabularies. It is a structured vocabulary developed primarily for the field of art history, but with the potential for wide applications in related disciplines such as archeology, history and geography. Looking at the case of the Thesaurus for Graphic Materials we see that neither classification provides an instinctive category where a KOS dealing with 'Graphic Materials' is easily located. Symbols in the ICC bring us closer to the subject matter of 'terms to index types of photographs, prints, design drawings, ephemera, and other pictures.' However the TGM poses an additional challenge in that the thesaurus does not deal with terms related to Graphic Materials but rather terms used for *annotation of* Graphic Materials. Terms included in this KOS range from 'Animal welfare' to 'Lost property'.

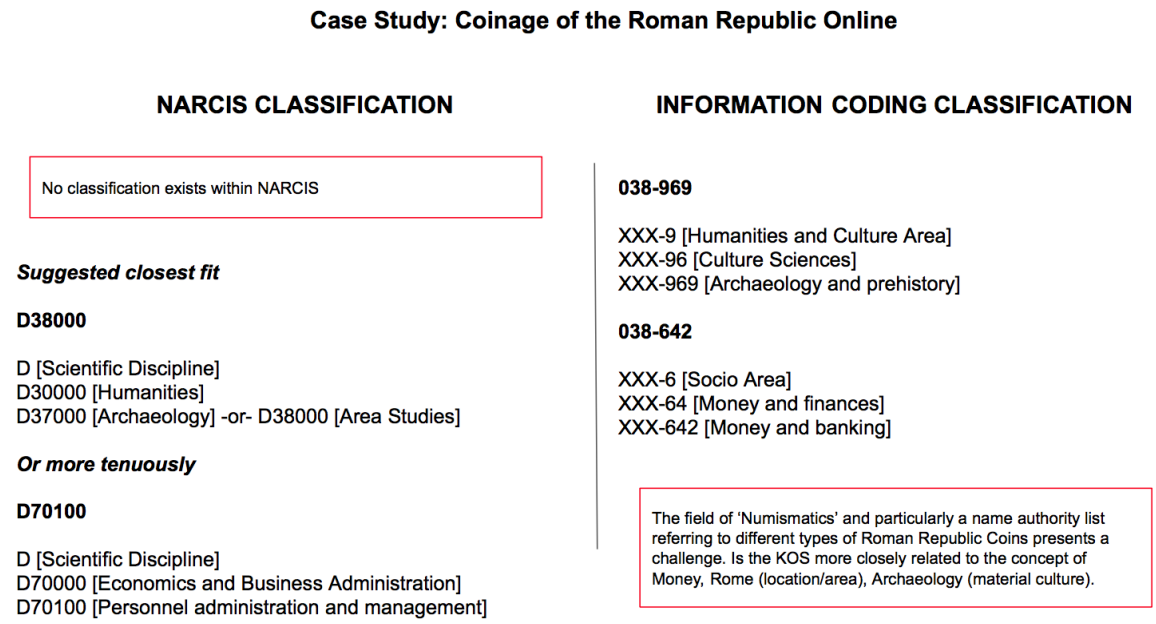

Figure 7. **Classification of Coinage of the Roman Republic Online.**

Coinage of the Roman Republic Online is a name authority list referring to different types of Roman Republic Coins. It there deals with terms in the field of 'Numismatics' and presents a particular challenge for both classifications. The classifier in this case must consider if the KOS more closely related to the concept of Money, Rome (Location/Area), or Archaeology (Material

Culture). As highlighted in Figure 7, for this KOS no classification exists within NARCIS although some suggestions have been made such as D37000 [Archaeology] or potentially D38000 [Area Studies].

5.0 Discussion and Conclusions

With this set of five new cases we add to the evaluation of classification of KOSo content reported in Coen and Smiraglia (2019). We are continually enlightened by the multi-dimensional flexibility of the ICC for classifying knowledge. In almost all cases we find hospitality in the ICC schedules for the fields represented in KOSo. We also are impressed with the accuracy of the NARCIS Classification, which, although a bit clunky in terms of classification structure, remains a viable and pragmatic representation of the fields under the umbrella of the Dutch national research community. Together, we believe the two classifications provide a rich landscape for future discovery when KOSo is able to go live and become searchable online.

More to the point, this latest exercise in comparative classification has demonstrated how the ICC's meta-ontological epistemic stance is particularly appropriate for an "observatory" of KOSs, the domains of which cross all forms of knowledge discourse. A simple classification such as NARCIS provides sufficient clarity for the pinpointing of scientific discourse, but ICC is a better multi-dimensional pointer to the ontological structuring of sciences within other realms of human knowledge. Using both together in the DANS KOS Observatory will provide users with both clarity of scientific positioning, on the one hand, and ontological relativity, on the other.

References

Beghtol, Clare. 2009. “Classification Theory.” In M. Bates & M. Niles Maack, *Encyclopedia of Library and Information Sciences*, 3rd ed. Boca Raton: CRC Press, 16.

Coen, Gerard and Richard P. Smiraglia. 2019. “Toward Better Interoperability of the NARCIS Classification.” *Knowledge Organization* forthcoming.

Dahlberg, Ingetraut. 1982a. “ICC--Information Coding Classification—Principles, Structure and Application Possibilities. *International Classification* 9: 87-93.

Dahlberg, Ingetraut. 1982b. “Annex 2: Information Coding Classification.” *International Classification and Indexing Bibliography* 1: 114-32.

Dahlberg, Ingetraut. 1999. “Classification System for Knowledge Organization Literature.” *Knowledge Organization* 26: 192-202.

Dahlberg, Ingetraut. 2008. “The Information Coding Classification (ICC): A Modern, Theory-Based Fully-Faceted, Universal System of Knowledge Fields.” *Axiomathes* 18: 161-76. doi:10.1007/s10516-007-9026-8

Fox, Melodie J. 2016. “Subjects in Doubt: The Ontogeny of Intersex in the Dewey Decimal Classification.” *Knowledge Organization* 43:581-93.

Lauruhn, Michael and Paul Groth. 2016. “Sources of Change for Modern Knowledge Organization Systems.” *Knowledge Organization* 43:622-29.

Lee, Wan-Chen. 2016. “An Exploratory Study of the Subject Ontogeny of Eugenics in the New Classification Scheme for Chinese Libraries and the Nippon Decimal Classification.” *Knowledge Organization* 43:594-608.

Mazzocchi, Fulvio. 2018. “Knowledge Organization System (KOS): An Introductory Critical Account.” Knowledge Organization 45(1): 54-78.

Szostak, Rick, Andrea Scharnhorst, Wouter Beek, Richard P. Smiraglia. 2018. “Connecting KOSs and the LOD Cloud.” In *Challenges and Opportunities for Knowledge Organization in the Digital Age: Proceedings of the Fifteenth International ISKO Conference, 9-11 July 2018, Porto, Portugal*, ed. Fernanda Ribeira and Maria Elisa Cerveira. Advances in Knowledge Organization 16. Baden-Baden: Ergon, 521-29.

Salah, Almila Akdag, Cheng Gao, Krzysztof Suchecki, Andrea Scharnhorst, and Richard P. Smiraglia. (2012). “The Evolution of Classification Systems: Ontogeny of the UDC.” In *Categories, Contexts and Relations in Knowledge Organization: Proceedings of the Twelfth International ISKO Conference 6-9 August 2012 Mysore, India* , ed. A. Neelameghan and K.S. Raghavan. Advances in Knowledge Organization 13. Würzburg: Ergon Verlag, 51–57.

Scharnhorst, A., C. Gao, A. Akdag Salah and K. Suchecki. 2012. “Evolution of Wikipedia Categories.” DANS. http://dx.doi.org/10.17026/dans-xjpzfuw

Scharnhorst, Andrea and Richard P. Smiraglia. 2012. “Evolution of Classification Systems.” *Advances In Classification Research Online* 23: 56. doi:10.7152/acro.v23i1.14264

Scharnhorst, Andrea, Richard P. Smiraglia, Christophe Guéret and Alkim Almila Akdag Salah. 2016. “Knowledge Maps of the UDC: Uses and Use Cases.” *Knowledge Organization* 43:641-54.

Smiraglia, Richard P., Andrea Scharnhorst, Almila Akdag Salah and Cheng Gao. 2013. “UDC in Action.” In Classification and Visualization: Interfaces to Knowledge, Proceedings of the International UDC Seminar, 24-25 October 2013, The Hague, The Netherlands, ed. Aïda Slavic, Almila Akdag Salah and Sylvie Davies. Würzburg: Ergon-Verlag, 259-72.

Tennis, Joseph. T. 2002. “Subject Ontogeny: Subject Access through Time and the Dimensionality of Classification.” In *Challenges in Knowledge Representation and Organization for the 21st Century: Integration of Knowledge across Boundaries: Proceedings of the Seventh International ISKO Conference, 10-13 July 2002, Granada, Spain*, ed. M. J. López-Huertas. Advances in Knowledge Organization 8. Würzburg: Ergon Verlag, 54-9.

Tennis, Joseph T. 2007. “Scheme Versioning in the Semantic Web.” *Cataloging & Classification Quarterly* 43, no.3: 85-104. DOI: 10.1300/J104v43n03

Tennis, Joseph T. 2012. “The Strange Case of Eugenics: A Subject’s Ontogeny in a Long-lived Classification Scheme and the Question of Collocative Integrity.” *Journal of the*

American Society for Information Science and Technology 63: 1350–59. doi:10.1002/asi.22686

Tennis, Joseph T. 2015. “The Memory of What Is: Ontogenic Analysis and its Relationship to Ontological Concerns in Knowledge Organization.” In *Ontology for Knowledge Organization*, ed. Richard P. Smiraglia and Hur-Li Lee. Würzburg: Ergon-Verlag, 97-100.

Tennis, Joseph T. 2016. “Methodological Challenges in Scheme Versioning and Subject Ontogeny Research.” *Knowledge Organization* 43: 573-80.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Bouwman, J. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data, 3.*

---

[1] Sponsored by COST Action grant TD1210. See Evolution and variation of classification systems – KnoweScape workshop March 4-5, 2015 Amsterdam (http://knowescape.org/evolution-and-variation-of-classification-systems-knowescape-workshop-march-4-5-2015-amsterdam/) and Workshop Observatory for Knowledge Organisation Systems in Malta, Feb 1-3, 2017 (http://knowescape.org/event/observatory-knowledge-organisation-systems/).