Observing trajectories of KOSs Across Space and Time: The DANS KOS Observatory (KOSo)

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

published in
Proceedings from North American Symposium on Knowledge Organization - NASKO
2019

DOI (link to publisher)
10.7152/nasko.v7i1.15619

document version
Publisher's PDF, also known as Version of record

document license
Unspecified

Gerard Coen, Richard P. Smiraglia, Peter Doorn **&** Andrea Scharnhorst. 2019. Observing Trajectories of KOSs Across Space and Time: The DANS KOS Observatory (KOSo). NASKO, Vol. 7. pp. 35-47.

Gerard Coen
Richard P. Smiraglia
Peter Doorn
Andrea Scharnhorst

Observing trajectories of KOSs Across Space and Time: The DANS KOS Observatory (KOSo)

Abstract

Knowledge Organization Systems (KOSs) include a wide variety of schemas ranging from ontologies, to classifications, thesauri, taxonomies, semantic networks, etc. These schemas can be updated and revised (or conversely become obsolete or lost) and are therefore prone to change over time. A wish expressed frequently by the research front in the KO community was for an "observatory" of KOSs. In 2017, via the KNAW Visiting Professor programme, DANS [1] began to focus more on understanding how KOSs change over time, how they can be archived, how version identification and control can be addressed, and also, how KOSs can be aligned to the 'FAIR' Data Principles (Findable, Accessible, Interoperable, Reusable). This research ambition coupled with community interest lead to the creation of the KOSo (Knowledge Organization Systems Observatory). Concretely, the observatory involves the identification of KOSs within the social sciences and humanities or the life sciences. KOSs have been described and ordered in the observatory through a process of empirical association in order to resist the potential pitfall of already organizing these resources through the lens of other KOSs (e.g. already describing the KOS in terms of existing controlled vocabularies). KOSo employs both metadata terms and formal classifications, using the Information Coding Classification in a synthetic format together with the KO Literature Classification, thus rendering for each KOS a domain-centric term faceted with a KOS-form term. Additionally, we classify domains using the NARCIS Classification, which is a framework to represent the research foci of the Dutch national research infrastructure.

"Do not look at stars as bright spots only. Try to take in the vastness of the universe."
— Maria Mitchell, Astronomer

Knowledge Organization Systems (KOSs) include a wide variety of schemas that organize, manage, and retrieve information and knowledge. They range from ontologies, to classification schemes, thesauri, taxonomies, controlled vocabularies and semantic networks (Mazzocchi, 2018). For organizations tasked with data stewardship, KOSs are crucial resources for the management and maintenance of data, permitting the ingestion of information from different sources through the application of schema mapping. These schemas can be updated and revised (or conversely become obsolete or lost) and are therefore prone to change over time. Schema mapping, also known as 'ontology alignment' within the field of Computer Science, is therefore a complex challenge exacerbated by issues with version identification and control. To cite an example provided by Scharnhorst (2015) "UDC numbers in MARC records (a data format for bibliographic information) almost never contain the number of the UDC edition from which they have been applied." The use of web technologies in scholarly communication and the emergence of Linked Data as specific data

format makes this time and version aspect of resources (including KOS) even more tangible (Klein et al., 2014).

The idea of a KOS observatory

At present the KOS landscape is a set of isolated, often domain-specific registries, with an array of disjunct purposes and scopes. It is also rapidly growing with the emergence of new research fields. Currently, we lack the possibility to step back and perceive the vastness of the KOS universe. For data archives, that means that data they are curating might not be interpretable in the future. But, also for Knowledge Organization (KO) scholars, appropriate empirical evidence is missing to explore the knowledge domain specificity, depth, age, and complexity of KOSs. Moreover, this research theme overlaps with the interest of scholars in the KO and Semantic Web (SW) communities. Szostak, Scharnhorst, Beek and Smiraglia (2018) wrote:

> A challenge frequently articulated across the KO and SW communities is the ability to track and maintain access to changing KOSs across time and across applications… Two KNOWeSCAPE workshops were held in Amsterdam (2015) and Malta (2017) to bring together experts from the KO, SW, publishing and digital humanities communities to prioritize objectives for visioning and creating an observatory for KOSs.

A wish expressed by the research front in the KO community is for an "observatory" of KOSs[1]. Many in the community have brought forward evidence of the need for a repository of KOSs that can allow observation across space and time. Tennis (2002; 2007; 2012; 2015; 2016) has been the most prominent catalyst for the KO community, joined by Fox (2016), Lauruhn and Groth (2016), Lee (2016), Salah et al. (2012), Scharnhorst et al. (2012; 2016), Scharnhorst and Smiraglia (2012) and Smiraglia et al. (2013). Such an observatory would be a place where one might find:

- All KOSs side by side;
- All instantiations of each such KOS side by side;
- Access to the classified worlds populated by instantiated KOSs (e.g., all UDC strings in WorldCat; all ICONCLASS terms in [some place]; etc.) side by side;

Provided with such a vantage point, one might, to use the metaphor of observatory literally, see the physical forces of specific designations of phenomena in cultural settings across time and across KOSs. In other words, an observatory that allows us to make observations, and capture observations, at once at many points in space-time.

[1] See the contributions to two workshops organised under the COST Action TD1210 (KNOWeSCAPE). First workshop on "Evolution and variation of classification systems" was held March 4-5, 2015 Amsterdam, http://knowescape.org/evolution-andvariation-of-classification-systems-knowescape-workshop-march-4-5-2015- amsterdam/ and the second was the Workshop on "Observatory for Knowledge Organisation Systems "held in Malta, Feb 1-3, 2017 (http://knowescape.org/event/observatory-knowledge-organisation-systems/).

As emphasized by Smiraglia (2017), an observatory allows observations at a point in time that capture at once many points in space-time. Each point has its own temporal reality. Cultural synergy (Smiraglia, 2014) tells us that each point in a KOS captures not just past knowledge, but also its cultural and social epistemology, an observatory, then, must capture not just the KOSs in their current states at given moments but also the cultural milieu attendant with each. Otherwise, the same thing at different times appears to be different, because it is different. Cultural synergy is the idea that information institutions (and KOSs are information institutions) can provide synergistic action through cultural interplay. This is based on the idea that institutions are disseminators of the cultures from which they spring. Therefore, we need, not just a repository of KOSs, nor a sequence of instantiations of KOSs. Rather, we need a network of interfacing cultural snapshots with temporal milestones. Cultural temporality, in other words, is the raison d'être for an observatory of KOSs. This reflection informs the set-up of the observatory, the attributes and provenance documented for each KOS, the workflow to maintain (update) and preserve long-term the observatory itself, and its content, the classification of classifications applied, aspects we detail below, and which define the specificity of the observatory compared to other initiatives to register KOSs.

Furthermore, research into better efficacy for the use of the LOD Cloud by researchers in the Social Sciences and Humanities (SSH) makes it clear that a repository and temporally functional observatory of all KOSs—from social classifications, to scientific taxonomies, to information architecture semantic web ontologies—is necessary, if only to track usage of controlled vocabularies in the Cloud.

The DANS KOS observatory - Overview of KOSo

DANS (Data Archiving and Networked Services) is the institute tasked with the stewardship of national research information in the Netherlands. It is an institute of 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). Its services include **NARCIS**, the Dutch national portal for research information, **EASY**, an online archiving system for depositing and reusing research data (Coen & Smiraglia, 2019), and **DATAVERSE**, for online storage, sharing and registration of research data, during the research period and up to ten years after its completion. Alongside these core services, DANS also participates in (inter)national projects and infrastructures, such as DARIAH (Digital Research Infrastructure for the Arts and Humanities), CESSDA (Consortium of European Social Science Data Archives), EOSC-Hub (European Open Science Cloud-Hub) and the Trans-Atlantic Platform "Digging into Data" Challenge, to name but a few. In this way DANS actively contributes to sustainable access to research data.

In 2017, Professor Richard P. Smiraglia joined DANS as a Visiting Professor with the aim of further development of the DANS services and to develop more insight into the history, evolution and mutual dependencies of the different domain-specific KOSs DANS sees itself confronted with. Together with Professor Smiraglia, DANS began to focus more on

understanding how KOSs change over time, how they can be archived, how version identification and control can be addressed, and also, how KOSs can be aligned to the 'FAIR' Data Principles (Findable, Accessible, Interoperable, Reusable) (Wilkinson et al., 2016).

This research ambition coupled with community interest lead to the creation of the KOSo (Knowledge Organization Systems observatory) at DANS at the end of 2017. Concretely the observatory involves the identification of KOS within the research fields of the Social Sciences and Humanities or Life Sciences, and their subsequent documentation across a range of attributes with particular emphasis on those attributes dealing with structure (space) and time. KOSs have been described and ordered in the observatory through a process of empirical association in order to resist the potential pitfall of already organizing these resources through the lens of other KOSs (e.g. already describing the KOS in terms of existing controlled vocabularies). The data generated is a first step towards providing a range of baseline statistics for KOSs. Overcoming questions of maintenance, the project hopes to lead to a registry to search across different existing KOSs, including records of all versions. Furthermore, the project provides information for the development of policy regarding the appropriate archival, maintenance, FAIRification and citation of KOSs.

From the earliest days of the observatory project we were concerned with a variety of techniques for using both metadata terms and formal classification symbols to enrich our data. The discussion led to the Information Coding Classification (ICC) as one of the most obvious and appropriate KOS. The ICC was first described by Dahlberg (1982) and clarified in Dahlberg (2008) as a classification of knowledge fields, instead of disciplines, organized according to the theory of integrative levels. According to Dahlberg the ICC is "fully-faceted" because its matrix-like structure is placed at the intersection of the nine integrative levels with a set of nine general form categories (e.g., theories, application of methods, etc.) which accomplishes systematic placement of knowledge fields. As an essential third dimension, Dahlberg described synthesis using "combinatory functions" (172) representing the expression of phase relationships. The elegance of the three-dimensional matrix representation of knowledge embedded in the ICC is derived from its essentially meta-physical origins. The richness of its structure suggests the richness of knowledge itself, thus making it prima facie the best choice for classifying KOSs themselves. 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. Additionally, as we contemplated positioning the KOSo data among DANS online resources we began to classify domains using the DANS NARCIS Classification. 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. Means and techniques for classifying KOSo are described more fully in Coen and Smiraglia (2019), and Coen, Smiraglia, Doorn and Scharnhorst (2019).

Contents of KOSo: How to Observe a KOS

Using metadata terms and formal classification symbols to enrich the data presents its own challenges meaning that some important questions needed to be addressed early on in the

KOSo setup. In order to assess which attributes were most relevant to describe the KOSs, we considered long-form questions and the research objectives of the project. This allowed us to move from a question like 'where can I find this KOS?' to the creation of the KOSo attribute columns 'Physical Location' and 'Online Location'. This process also aided in narrowing down which data we should focus on, since there are not data available for answering all questions of interest in an objective manner. The question 'who are the users of this KOS?' while highly relevant remains challenging to answer in most circumstances and therefore no corresponding 'Users' attribute column exists.

Once it had been loosely decided which attributes of KOS should be investigated two further questions arose a) which metadata terms should the project adopt? b) should we already use controlled vocabularies in the KOSo to make the data easier to analyze and compare? Using Dublin Core metadata terms, or other schemas for the attributes and describing KOS using existing controlled vocabularies was largely avoided. The benefit of this was that the KOSs could be described in a way which was not limited by viewing them through a pre-existing lens, designed for another data type. The main disadvantage of this is that the data produced, true to the KOSs themselves, are quite heterogeneous.

Since the KOSo does not stand alone in its attempts to create a registry and classification of KOSs, efforts have been made to link to, collaborate with and incorporate aspects of other relevant initiatives. The Dublin Core NKOS group 'KOS Types Vocabulary' (http://dublincore.org/groups/nkos/) has been included as an attribute. Furthermore, when a KOS appears in both the KOSo and the Basel Register of Thesauri, Ontologies and Classifications [BARTOC] (https://bartoc.org), a URL is provided.

Currently the KOSo collects information about KOSs under twenty-three attributes as seen in Table 1:

Table 1. Attribute names and descriptions

Attribute Name	Attribute description
Identifier	A unique identifier given to each instantiation of a KOS.
Schema Name/Title	The name of the KOS including any aliases
Creator(s)/Curator(s)	The name of the original creator and the names of subsequent curators responsible for its maintenance.
Maintenance organization	The name of the institution responsible for publishing/maintaining the KOS.
Place: Publisher, date.	Where the KOS was published, the publisher and publication date.
Summary / Abstract	A description of the purpose of the KOS.
Format(s)	A list of the formats the KOS is published in.
Language	A list of the languages the KOS is published in.
Physical Location	Where a physical copy of the KOS can be found within the Netherlands i.e. a library address.
Online Location	Where the KOS can be found online i.e. a URL.
Earlier versions (editions) …	The previous version to this instance of the KOS.
History of versioning:	The version number/edition of the KOS.
Version Notes:	What has been changed, deleted or added since the previous version. Is there something else noteworthy about this version.
Last Updated:	The date of the last update of the KOS.
Number of terms in system	The number of terms the KOS includes.
Phenomena included	The *things* the KOS aims to organize/manage.
Disciplines included	The scientific disciplines covered by the KOS.
Direct domains included	The research domains, fields or themes covered by the KOS.
Related to:	Does the KOS reference any other KOS? It is branched from, merged with or structured in the form of another KOS.
BARTOC link	Does this KOS appear in the Basel Register of Thesauri, Ontologies and Classifications? If so, a URL is provided.
International Coding Classification (ICC)	The relevant ICC code for the KOS.
NARCIS Classification	The relevant NARCIS class for the KOS.
KOS Types Vocabulary	The appropriate Dublin Core NKOS group 'KOS Types Vocabulary' label for the KOS.

The data are collected in an excel file with twenty-three columns. Due to the size of the spreadsheets in the KOSo viewing nested inside a publication is not ideal, however the following figures provide an example of how the data are presented.

Figure 1. 'Identifier', 'Schema Name/Title' and 'Phenomena included' from the SSH KOSo

Identifier	Schema Name/Title	Phenomena included
KOS.HSS.35	MDA archaeological objects thesaurus	For the recording of portable archaeological objects.
KOS.HSS.35.1	Archaeological Objects Thesaurus	For the recording of portable archaeological objects.
KOS.HSS.36	Archaeological Sciences Thesaurus	Terminology used for recording the techniques, recovery methods and materials associated with archaeological sciences (e.g. tree-ring analysis, modification state, pathology)
KOS.HSS.37	Thesaurus of building materials : a standard for use in architectural and archaeological records	Includes natural/animal and man-made material (e.g. cement mix, fossil, fibre-glass).
KOS.HSS.37.1	Building Materials Thesaurus	Includes natural/animal and man-made material (e.g. cement mix, fossil, fibre-glass).
KOS.HSS.38	Cargo Thesaurus	Thesaurus for types of cargo carried by vessels and aircraft on their final voyage (e.g. animal fodder, sugar, fuel).
KOS.HSS.39	Components Thesaurus	Terminology covering divisions and structural elements of a building or monument. Includes terms that describe areas and spaces, decorative features, fixtures and fittings, machinery and implied features (e.g. post hole, cartouche, truss).
KOS.HSS.40	Defence of Britain Thesaurus	Terminology relating to the Second World War defence of Britain.
KOS.HSS.41	Event Thesaurus	Use for recording archaeological events, e.g. architectural investigative, data collection exercises; from intrusive interventions into the resource to non-damaging survey events (e.g. grab sampling, auger survey, remote sensing).
KOS.HSS.42	Evidence Thesaurus	Terminology covering the existing physical remains of a monument, or the means by which a monument has been identified where no physical remains exist (e.g. placename evidence, soilmark, surface deposit).
KOS.HSS.43	Farmstead Thesaurus	Thesaurus for indexing different types of farmsteads, related buildings, areas and layouts.
KOS.HSS.44	First World War Thesaurus	For the recording of monuments related to the First World War in the UK.
KOS.HSS.45	Heritage Crime Thesaurus	Thesaurus for indexing types of crime and incidents against heritage assets (e.g. graffiti, theft, unauthorized diving).
KOS.HSS.46	Historic Aircraft Thesaurus	Indexing terminology for the recording of aircraft remains and crash sites, listing aircraft types by form, function and manufacturer (e.g. Atlas, Target, Westland).
KOS.HSS.47	Historic Characterisation Thesaurus	Thesaurus combining Historic Landscape and Historic Seascape Characterisation terminology (e.g. reclaimed land, vineyard, dunes).
KOS.HSS.48	Manner of Loss List	A small authority file/ wordlist for indexing types of loss of vessels or aircraft (e.g. beached, lost, enemy action).
KOS.HSS.49	Maritime Craft Thesaurus	Craft types which survive as wrecks for Historic England's maritime record and can be used to describe types of ship (e.g. block ship, rescue vessel, net layer) .

Figure 1 shows an extract from the SSH KOSo highlighting the attributes 'Identifier', 'Schema Name/Title' and 'Phenomena included'. The information provided under 'Identifier' has been created by the researcher for the purpose of internal control. 'Schema Name/Title' was transcribed from the source, synonyms and abbreviations are also noted here. The information for the attribute 'Phenomena included' is generally derived from the summary or abstract explaining why the KOS was developed or its aim.

Figure 2. 'Creator(s)/Curator(s)', 'Maintenance Organization' and 'Place: Publisher, date' data

Creator(s)/Curator(s)	Maintenance organization	Place: Publisher, date.
Frau Prof. Hagedorn-Saupe	Institut für Museumsforschung (& Getty Art History Information Program)	Berlin : Institut für Museumsforschung, 2014
n/a	El Centro de Documentación de Bienes Pa (& Getty Art History Information Program)	Santiago : El Centro de Documentación de Bien
Henri van de Waal, Rudi Fuchs, Leendert Couprie	Koninklijke Nederlandse Akademie van We	Amsterdam : North-Holland, 1973-1981.
Henri van de Waal	Utrecht University	Leiden/Utrecht : ICONCLASS Research & Deve http://www.iconclass.nl/about-iconclass/history-c
Simona Ciofetta, Marco Lattanzi, Elena Plances,	Istituto centrale per il catalogo e la docume	Roma : Istituto Centrale per il Catalogo e la Doc
Reem Weda	RKD-Nederlands Instituut voor Kunstgesch	The Hague : RKD- Nederlands Instituut voor Ku
Robert G. Chenhall	n/a	Nashville : American Association for State and L
Robert G Chenhall; James R Blackaby; Patricia Greeno;	Nomenclature Committee, the.	Nashville : AASLH Press, 1989.
James R Blackaby; Patricia Greeno; Robert G Chenhall	Nomenclature Committee, the.	Walnut Creek, Calif. : AltaMira Press, 1995
Paul Bourcier; Ruby Rogers;	Nomenclature Committee, the.	Lanham, Maryland : Altamira Press, cop. 2010.
Paul Bourcier, Heather Dunn;	Nomenclature Task Force	New York : Rowman & Littlefield, 2015
College Art Association (U.S)	Getty Art History Information Program	Santa Monica, CA : The Getty Art History Inform [Los Angeles?] : College Art Association, 1995.
College Art Association (U.S)	Getty Art History Information Program	Santa Monica, CA : The Getty Art History Inform [?] : College Art Association, 1996.
Murtha Baca; Patricia Harpring;	The J. Paul Getty Trust	Santa Monica, CA : J. Paul Getty Trust ; [?] : College Art Association, ©2000.
Murtha Baca; Patricia Harpring;	Getty Art History Information Program, Coll	[Santa Monica, California] : J. Paul Getty Trust ; [?] : College Art Association, 2006-
Murtha Baca; Patricia Harpring;	Getty Art History Information Program	The J. Paul Getty Trust
J. D. Wees	Library of Congress	VRA BULLETIN, 23, no. 3, (1996): 57-58
VRA Data Standards Committee	Library of Congress	Washington : VRA Data Standards Committee, 2
Alice Grant; Joséphine Nieuwenhuis; Toni Petersen;	International Committee for Documentation	Paris : CIDOC, 1995.

Figure 2, shows an example of information contained in the fields 'Creator(s)/Curator(s)', 'Maintenance Organization' and 'Place: Publisher, date'. Initially the 'Creator(s)/Curator(s)' column was labelled 'Creator' but was later changed to accommodate the fact that KOSs can have many contributors over time. This information was either taken from bibliographic records such as those in the OCLC WorldCat or from Editor/Contact information provided within the KOS. 'Maintenance Organization' references the institute within which the creators/curators work, or in some cases who is responsible for funding the development and the upkeep of the system. This information can be found in bibliographic records, but also in the 'About' section of a KOS published online. In some cases, for systems online, the maintenance organization can be seen in the URL pathway, or at the bottom of a webpage in the form of copyright, privacy and disclaimer information. The attribute 'Place: Publisher, date' is also most often transcribed from bibliographic information sourced from WorldCat.

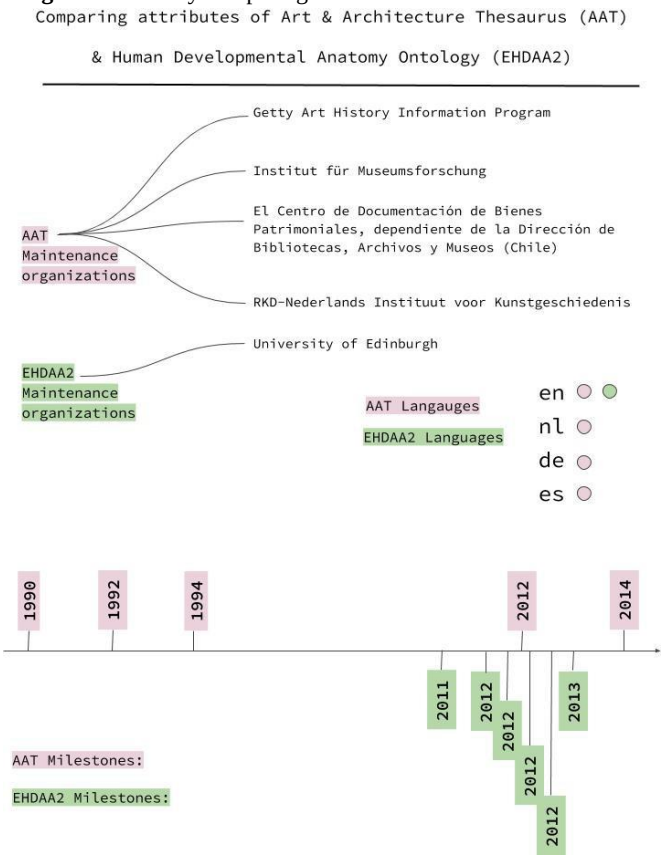
Figure 3. 'Earlier versions (editions)', 'History of versioning' and 'Version Notes' data

Earlier versions (editions) ...	History of versioning:	Version Notes:
n/a	Version 0.9 B	"This beta release is a draft for your review and evaluation before its final release in February 1996"--p. [3] of Quick reference card.
Version 0.9 B	Version 1.0	Accompanies special issue of Visual resources, v. 11, no. 3-4, 1996.
Version 1.0	1st Edition	n/a
1st Edition	Online Version	"Revised March 2014 by Patricia Harpring." Earlier versions issued in print.
Online Version	Revised 2016	Revised 2016 by Patricia Harpring. Murtha Baca and Patricia Harpring, Editors
n/a	Version 1.0	Originally published as Wees, Dustin J., "Core Categories for Visual Resources. A Draft Proposed by the VRA Data Standards Committee." VRA Bulletin 23, 3 (Fall 1996), pp. 57-63.
Version 3.0	Version 4.0 Core Schemas; Unrestricted Version	The current version 4.0, that was released in 2007, is expressed as an XML schema in order to support the interoperability and exchange of VRA Core records. The unrestricted schema imposes no requirements on the values entered into any of the elements, sub-elements, or attributes, and may be useful for those who want to exchange legacy data.
n/a	Version 1	http://old.cidoc-crm.org/docs/guidelines/guidepre.htm
Version 6.2.1	Version 6.2.2	Editorial Status: Open-In progress
Darwin Core 1.4 (Draft Standard)	Date Issued: 12/02/2009	A TDWG task group was created to revise the Darwin Core, and a ratified metadata stan Before becoming a TDWG standard, the Darwin Core was instantiated and deployed in n
n/a	2010-11-08: version 1.0.	http://network.icom.museum/cidoc/working-groups/lido/lido-overview/lidos-background/
Version 1.0	Printed Book Version	WP3 working group "Identifying standards and developing recommendations".
1st Edition (1994)	2nd Edition (1997)	n/a
n/a	Version 5.0	n/a

Figure 3, provides an example of the type of information seen in the columns 'Earlier versions (editions)', 'History of versioning' and 'Version Notes'. Data identifying 'Earlier versions (editions)' is provided when the immediately preceding version of that being described is known. 'History of versioning' details which version is being described; this information is most often contained in a release history but can also be deduced by searching for all versions of the KOS and placing them in chronological order. The 'Version Notes' are provided when some comment has been discovered about what has changed or been updated since the prior version. For example, in the case of KOSs published on Github this information has been taken from the 'Commit History'.

While not all data allow for a very meaningful visual comparison, some interesting insights can still be gained by placing certain attributes side-by-side. Below we explore a brief comparison using one example each from the SSH & LS KOSo.

Figure 4. Visually comparing some attributes of AAT & EHDAA2



At a glance Figure 4 tells us that the AAT KOS is older than the EHDAA2 but also it is available in more languages and has had more maintenance organizations. While both KOS have five versions, those of the AAT happened over a longer period and with more gaps in between while EHDAA2 had a burst of activity between 2011 and 2013 and has not seen much activity since. Viewing this small case study, we can already imagine with more data which types of insights and inferences could be made which would be of interest both to scholars of KO and also the emerging field of Science of Science (SciSci) (Scharnhorst et

al., 2012). SciSci "explores patterns characterizing the structure and evolution of science" (Fortunato et al., 2018, p.2)

FAIR and a Public Portal for KOSo

Recognising the value of current and previous versions of KOSs in providing both the access to and context for interpreting data, in its 2018-2020 research programme DANS has set out an objective to become an archive for "Endangered Knowledge Organisation Systems" (Aerts et al., 2018). The first step in this journey has been to gain more familiarity with KOSs as part of the KOSo project. The next step in achieving this ambition has been to consider technically how to make the data collected in the KOSo project Open and FAIR so that it becomes available to the wider KO and SW communities. This process involves learning about the handling of these resources and their associated metadata from an archival standpoint.

Making KOSs themselves FAIR involves their classification and description using appropriate rich metadata, also ensuring that the data are published in a sustainable environment. This involves, defining the metadata needed to make KOSs FAIR and also establishing or uncovering the best practices for the citation of KOSs. Currently research is underway to evaluate how to best expose the KOSo within DANS' existing portfolio of services, namely NARCIS and Dataverse (https://dataverse.nl). The medium-term ambition of these efforts is to develop a public portal for KOSs.

Conclusion: A real Observatory

Our use of the "observatory" metaphor is intentional; the vision of a locus for visualization and dynamic observation of KOSs across spatial and temporal trajectories is both tantalizing for research and at the same time provides functional imperatives. Earlier work to survey KOSs at a meta-level (e.g., BARTOC or resources (such as those provided by the Dublin Core NKOS Interest Group), have developed into rich repositories. We hoped to add dimensionality in several ways—primarily by providing from the beginning an overview of instantiated KOSs, but also in future (and more ambitiously) by providing portals into all of the instantiations of a KOS. Thus, as we noted earlier, we decided not only to classify our data, but also to use overlapping classifications and metadata. Thus, we employ Dahlberg's Information Coding Classification synthesized for use with KOSs by combination with her KO Literature Classification. We also use the DANS-supported NARCIS Classification, to integrate our work with the DANS infrastructure. Interestingly (see Coen and Smiraglia 2019 and Coen, Smiraglia, Scharnhorst and Doorn 2019), the two classifications overlap at the level of disciplines but provide distinctly different ontological coverage at a more granular level. Combined with the richness of natural language terms derived from the KOSs themselves we have been able to overlay the KOSo with a rich ontological matrix that provides opportunities for comparative resource discovery that can provide Wilsonian exploitative power (Wilson 1968) for KOS research. The multi-pronged approach also helps us overcome the latent gaps or historical gaffes in both ICC and NARCIS. To carry the observatory metaphor one step further, we envision KOSo functioning much like an archives

— a portal for independent research, the products of which eventually can further enrich the observatory itself with highly precise domain ontologies.

References

Aerts, P., Coen, G., Dillo, I., Doorn, P., Morselli, F., & Scharnhorst, A. 2018. DANS Research Programme 2018-2020. doi: 10.17026/dans-xpc-abaj

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

Coen, Gerard, Richard P. Smiraglia, Peter Doorn and Andrea Scharnhorst. 2019. Classifying KOSs: The Use of Dahlberg's ICC in the DANS KOS Observatory (KOSo). Submitted to *ISKO-Low Countries.*

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

Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., ... & Vespignani, A. 2018. Science of science. *Science* 359: 6379.

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

Klein, Martin, Herbert Van de Sompel, Robert Sanderson, Harihar Shankar, Luydmila Balakireva, Ke Zhou, and Richard Tobin. 2014. Scholarly Context Not Found: One in Five Articles Suffers from Reference Rot. *PLoS ONE*, 9(12): e115253. doi:10.1371/journal.pone.0115253 http://dx.doi.org/10.1371/journal.pone.0115253

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, F. 2018. Knowledge organization system (KOS): an introductory critical account. *Knowledge Organization* 45: 54-78.

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, Andrea. 2015. After Evolution and Variation of Classification Systems- a workshop report by Andrea Scharnhorst - KNOWeSCAPE. Retrieved from http://knowescape.org/after-evolution-and-variation-of-classification-systems-a- workshop-report-by-andrea-scharnhorst/

Scharnhorst, Andrea, Gao, Cheng, Salah, Almila Akdag and Krzystof 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.

Scharnhorst, Andrea, Katy Börner and Peter van den Besselaar, eds. 2012. *Models of Science Dynamics: Encounters between Complexity Theory and Information Sciences*. Berlin; Heidelberg: Springer.

Smiraglia, Richard P. 2014. *Cultural Synergy in Information Institutions*. New York: Springer.

Smiraglia, Richard P. 2017. KOS on a Space-Time Continuum: Empirical Implications for a KOS Observatory. Keynote presentation at KnoweScape Workshop--Observatory for Knowledge Organisation Systems in Malta, Feb 1- 3, 2017. (http://knowescape.org/event/observatory-knowledge-organisation- systems/)

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.

Szostak, Rick, Andrea Scharnhorst, Wouter Beek and 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.

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.

Wilson, Patrick. [1968] 1978. *Two Kinds of Power: An Essay in Bibliographical Control*. California Library Reprint Series ed. Berkeley: University of California Press.