

Contents

Special Issue: Research Information Systems and Science Classifications; including papers from “Trajectories for Research:

Fathoming the Promise of the NARCIS Classification,” 27-28 September 2018, The Hague, The Netherlands.

Guest editors: Andrea Scharnhorst and Peter Doorn

Articles

Richard P. Smiraglia.
Trajectories for Research: Fathoming the Promise
of the NARCIS Classification..... 337

Gerard Coen and Richard P. Smiraglia.
Toward Better Interoperability of the NARCIS
Classification 345

Sadia Vancauwenbergh and Hanne Poelmans.
The Flemish Research Discipline Classification
Standard: A Practical Approach 354

Shenghui Wang and Rob Koopman.
Embed First, Then Predict..... 364

Reviews of Concepts in Knowledge Organization

Ariadne Legendere.
The Development of the Canadian Research
and Development Classification..... 371

Brief Communication

Rahul Thorat and Reinout van Brakel.
Brief Communication: The Need for a
National-Level Working Group for Higher
Education Research Data in The Netherlands380

Roman Solc.
Brief Communication: Some Critical Notes on the
Czech System of Evaluation of Research387

Letter to the Editor

Guohua Xiao.
Explainable AI: Linking Human and Machine398

Books Recently Published.....400

Contents pages

Smiraglia, Richard P. 2019. "Trajectories for Research: Fathoming the Promise of the NARCIS Classification." *Knowledge Organization* 46(5): 337-344. 13 references. DOI:10.5771/0943-7444-2019-5-337.

Abstract: NARCIS—National Academic Research and Collaborations Information System—is the national research portal for the Netherlands' data and research archiving, which is governed by its own NARCIS Classification. The current instantiation of the classification dates from 2015. The classification is currently made up of two classes—D for the sciences broadly, and E for interdisciplinary areas. The NARCIS Classification is designed specifically and with care for the contents of the NARCIS data portal. The classification mostly represents the sciences. A few anomalous situations are visible in the ontology of the classification: the humanities occupy one division within the sciences, placed between the life sciences and law; and, the treatment of interdisciplinarity, for which a separate class E is set aside for interdisciplinary sciences. A dump of the NARCIS database was used to analyze the population of the NARCIS classification. The life sciences occupy 34% of the NARCIS database. A framework for research networking systems reveals the NARCIS database and its classification meet most objectives, with the only lapse being the output of entities and attributes to ontologies. The NARCIS Classification is also an occupational classification. The NARCIS Classification supports a vital research portal that, in turn, supports a nationally-coordinated research effort designed to provide better inter-institutional communication of scholarly productivity, thus is in itself an information institution, in which domain-dependence is part of its cultural imperative. The NARCIS Classification incorporates an example of top-down politics in which funded disciplines are included and best represented. A perhaps unintended consequence is the encapsulation of forced views. Trajectories for further discussion with regard to continued development of the NARCIS Classification include identity, interoperability, interdisciplinarity, and synthesis.

Coen, Gerard and Richard P. Smiraglia. 2019. "Toward Better Interoperability of the NARCIS Classification." *Knowledge Organization* 46(5): 345-353. 9 references. DOI:10.5771/0943-7444-2019-5-345.

Abstract: Research information can be useful to science stakeholders for discovering, evaluating and planning research activities. In the Netherlands, the institute tasked with the stewardship of national research information is DANS (Data Archiving and Networked Services). DANS is the home of NARCIS, the national portal for research information, which uses a similarly named national research classification. The NARCIS Classification assigns symbols to represent the knowledge-bases of contributing scholars. A recent research stream in knowledge organ-

ization known as comparative classification uses two or more classifications experimentally to generate empirical evidence about coverage of conceptual content, population of the classes, and economy of classification. This paper builds on that research in order to further understand the comparative impact of the NARCIS Classification alongside a classification designed specifically for information resources. Our six cases come from the DANS project Knowledge Organization System Observatory (KOSo), which itself is classified using the Information Coding Classification (ICC) created in 1982 by Ingetraut Dahlberg. ICC is considered to have the merits of universality, faceting, and a top-down approach. Results are exploratory, indicating that both classifications provide fairly precise coverage. The inflexibility of the NARCIS Classification makes it difficult to express complex concepts. The meta-ontological, epistemic stance of the ICC is apparent in all aspects of this study. Using the two together in the DANS KOS Observatory will provide users with both clarity of scientific positioning and ontological relativity.

Vancauwenbergh, Sadia and Hanne Poelmans. 2019. "The Flemish Research Discipline Classification Standard: A Practical Approach." *Knowledge Organization* 46(5): 354-363. 14 references. DOI:10.5771/0943-7444-2019-5-354.

Abstract: In 2010, a study was performed by the Flemish universities in cooperation with the Flemish Interuniversity Council (VLIR) on the administrative burden of research reporting in Flanders, Belgium. One of the most prominent observations of this study (Peters and Lambrechts 2011) consisted of the redundancy that occurs both in preserving, classifying and reporting research information to different stakeholders in a region as small as Flanders. In response to this study, the Flemish government assigned the Centre for Research & Development Monitoring (ECCOM) with the task to: 1) develop a research discipline classification standard for the Flemish region that could serve all existing use purposes; 2) effectuate the implementation of this research classification standard by all Flemish stakeholders; and, 3) prevent data loss when classification schemes would be converted. This paper discusses the background, creation and implementation of the Flemish Research Discipline Classification Standard.

Wang, Shenghui and Rob Koopman. 2019. "Embed First, Then Predict." *Knowledge Organization* 46(5): 364-370. 13 references. DOI:10.5771/0943-7444-2019-5-364.

Abstract: Automatic subject prediction is a desirable feature for modern digital library systems, as manual indexing can no longer cope with the rapid growth of digital collections. It is also desirable to be able to identify a small set of entities (e.g., authors,

citations, bibliographic records) which are most relevant to a query. This gets more difficult when the amount of data increases dramatically. Data sparsity and model scalability are the major challenges to solving this type of extreme multi-label classification problem automatically. In this paper, we propose to address this problem in two steps: we first embed different types of entities into the same semantic space, where similarity could be computed easily; second, we propose a novel non-parametric method to identify the most relevant entities in addition to direct semantic similarities. We show how effectively this approach predicts even very specialised subjects, which are associated with few documents in the training set and are more problematic for a classifier.

Legendre, Ariadne. 2019. "The Development of the Canadian Research and Development Classification." *Knowledge Organization* 46(5): 371-379. 25 references. DOI:10.5771/0943-7444-2019-5-371.

Abstract: I describe the development of a new research and development taxonomy to facilitate the reporting of granting agency investments in research and the organization of effective peer review processes in Canada, which represents a kind of classification designed to support the administrative management of research. The development of the Canadian Research and Development Classification (CRDC) is being led by the Canada Foundation for Innovation (CFI), the Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Social Sciences and Humanities Research Council of Canada (SSHRC), in close collaboration with Statistics Canada. This collaboration represents an unprecedented effort by the federal research funding agencies and Statistics Canada to develop a common standard for the classification of research and development activities in Canada, and is intended to meet different needs within the broader research ecosystem. The CRDC, intended to align the way research and development is categorized in Canada, is the result of months of reviews, consultations, analysis and negotiations among the agen-

cies and the Canadian research community. Notably, the CRDC was designed to include all sectors of research and development, represent the current research landscape in Canada, support a wide range of needs within the research and development ecosystem, and increase computability with international standards.

Thorat, Rahul and Reinout van Brakel. 2019. "Brief Communication: The Need for a National-Level Working Group for Higher Education Research Data in The Netherlands." *Knowledge Organization* 46(5): 380-386. 2 references. DOI:10.5771/0943-7444-2019-5-380.

Abstract: We describe the struggle the Association of Dutch Universities (VSNU) faces to get proper data on research systems. The type of users the VSNU encounters (from universities to policy makers) and what that means for classifying research is described. We list the research data governance practices from various countries. Based on those practices, a working group involving various stakeholders to develop a common research data governance framework in the Netherlands is proposed. The working group would propose standards in higher education, research and impact specific to data collection, metadata and their interoperability across various stakeholders. The results of the first brainstorming session on the governance of research data are mentioned.

Solc, Roman. 2019. "Brief Communication: Some Critical Notes on the Czech System of Evaluation of Research." *Knowledge Organization* 46(5): 387-397. 4 references. DOI:10.5771/0943-7444-2019-5-387.

Abstract: We try to critically analyze some aspects of the system of evaluation of scientific research in the Czech Republic. The government presents it as very objective, fair and motivating. Using the evaluating methodology of the state, data obtained from *Journal Citation Reports* and mathematical models we would like to demonstrate that the system is neither fair nor motivating.