380

Knowl. Org. 46(2019)No.5
Brief Communication

Brief Communication:

The Need for a National-Level Working Group for

Higher Education Research Data in The Netherlands

Rahul Thorat*, Reinout van Brakel**

Association of Universities in the Netherlands (VSNU),
1 Lange Houtstraat 2 Den Haag, The Netherlands,
*<rahul@thorat.net>, **<vanbrakel@vsnu.nl>

Rahul Thorat (1980) is an institutional researcher and data scientist at the VSNU until 1 July 2019. He graduated in mechanical engineering in India and obtained his PhD in civil engineering and technical geosciences at TU Delft. Before Rahul joined the VSNU, he was a senior project officer (training and projects) at Data Archiving and Networked Services (DANS) -KNAW. At the VSNU, he works primarily for the development of higher education sector-wide key figures and a sector overview. He also makes a constructive contribution to cohesion among the members of the association.

Reinout van Brakel is Domain Leader for the Governance and Accountability section of the VSNU. Together with his team, he ensures that opinions about the Dutch university education and research can be substantiated with facts. He also ensures that the information is made public by developing fact sheets and infographics. Reinout is also involved in developments in the field of governance, legislation and regulations regarding management and supervision and the various codes of conduct that are used within the higher education sector. Prior to joining the VSNU, Reinout was a policy researcher and advisor at Ecorys and PwC.

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.

1.0 Introduction

About the Association of Research Universities in the Netherlands (VSNU): the history of VSNU goes back to its original idea in the 1940s. After initial "Interuniversitair Contactorgaan" (1956) and the Academic Council (1960), the Association of research universities in the Netherlands acquired its present form in 1985 with greater autonomy for the universities. The VSNU is partly comparable to such organizations from other countries (e.g., the German Hochschulrektorenkonferenz). The VSNU visits regularly its counterpart organizations from the other countries and receives those organizations at its location in The Hague. More details can be found on the VSNU website (https://www.vsnu.nl/nl_NL).

By joining forces through VSNU, research universities show the outside world how they fulfill their societal task, that is, to provide high-quality academic education and to conduct high-quality scientific research in order to build a strong knowledge society. The universities formulate common goals for education and research and lobby for the conditions that are necessary to achieve these goals.

Knowl. Org. 46(2019)No.5
Brief Communication

381

The VSNU office in The Hague supports these goals by providing a platform for knowledge exchange, by informing universities about the political realities, by lobbying towards national politics and by fulfilling the role of employer. Lobbying for conditions such as fewer rules and more funding is more successful when universities make their ambitions and successes visible and when they are accountable for their choices and actions.

VSNU exchanges data with various sources (Figure 1), which include universities, semi-government institutes, private parties and individuals. The collected data are published online via VSNU "Facts & Figures" pages. The objective of data collection and visualization is to offer transparency and accountability of the fourteen universities of the Netherlands to the government and to the public. Further, the VSNU has agreed with the Ministry of Education, Culture and Science (OCW)[2] to publish a national level overview of strategic information on university education, research and its impact on the Dutch society. A first version of this sector overview was published on 28 March 2019 on the VSNU website (the link to sector overview is given in the appendix).

1.1 Fragmented research data governance

The background data of the overview need to be updated from multiple data sources. In its current format, the data collection, update and visualization from multiple sources is a tedious task and is beyond the limited capacity of the VSNU. A collective research information (CRIS) system for research data as a joint venture with other stakeholders (universities, semi-government, government institutes) would ease the task. The stakeholders, involved in governing the information on research output, are individual universities, national-international organizations, expert working groups, ranking agencies, government institutions and research companies. There is no consensus on the classification among the aforementioned stakeholders to ensure that the information is relevant for policymakers and used in decision making on the national level politics. Further disadvantages of fragmented research data governance lead to labor-intensiveness, and to unclear legal and operational status of research data.

Figure 1 shows the current complex situation concerning the collection and exchange of data about research. We have listed all important partners and organizations (a list explaining the abbreviations can be found in the appendix). The arrows, colours and icons indicate information flows and their current status. The pink colored arrows represent data flowing to the VSNU from partner organizations. The blue colored arrows represent data flowing from the VSNU to partner organizations. The magenta colored arrows represent indicators and the yellow colored arrows represent services from private parties (e.g., ranking agen-
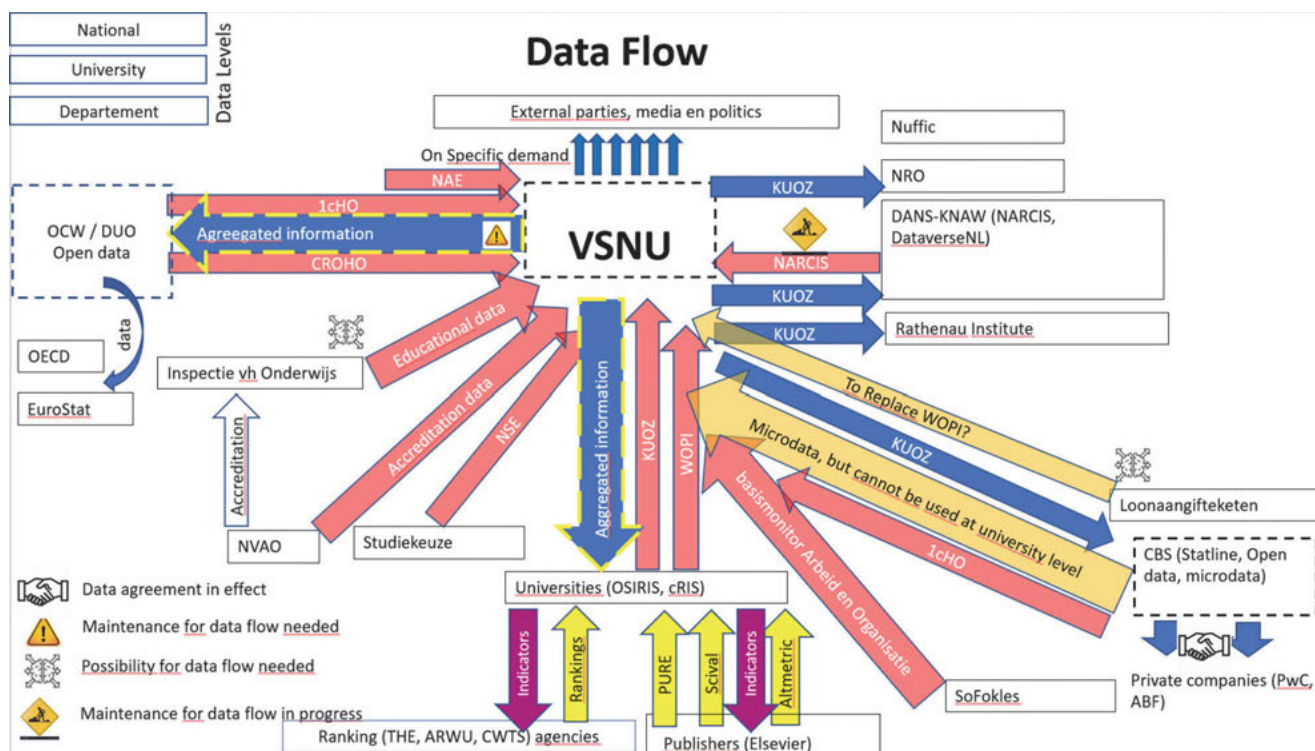

Figure 1. Current research data landscape for higher education system in The Netherlands.

382

Knowl. Org. 46(2019)No.5
Brief Communication

cies, data collecting/managing/publishing companies). The dotted borders of arrows and text boxes represent transparency towards data exchange. The full names behind all the abbreviations and the rough English translation of Dutch names are given in table 1. The situation of every data exchange is explained, where the data agreement is reached, where there is possibility of data agreement, where the maintenance for data exchange is ongoing and where the maintenance for data exchange is needed. There are also various levels of data aggregation: national levels, university levels and department levels.

1.2 Case study: problem of counting publications via various RISs

The fragmented governance leads to differences in counting research publications from the universities. Data on research input and output have been collected by the VSNU for many years and are known as KUOZ data (Kengetallen Universitair Onderzoek). The VSNU collects and manages the database, including data such as the number of publications, theses etc., based on classification developed by expert groups from the universities.

The information is processed and is shared with universities, press and with the OCW. However, KUOZ is one of many databases used for recording research data from the universities. Data Archiving and Networked Systems (DANS) operates a service National Academic Research and Collaborations Information System (NARCIS) by harvesting data from the universities and available repositories. The goals of both KUOZ and NARCIS are similar, to provide research data information to users. However, there are differences in counting publications as shown in Figure 2.

The difference in counting publications could be due to (table 2): 1) difference in registration of publications—not all university researchers and their publications are registered in NARCIS, while VSNU receives yearly publication numbers from the universities; 2) different ways of counting—KUOZ considers publications per university while NARCIS considers publications per authors from different universities; further, NARCIS may have all open access publications, but not all closed access publications are available; or, 3) difference in classification definitions and aggregation level—the definitions of scientific publications used in NARCIS and KUOZ varies.

Abbreviation	Name	Description/ English translation
1cHO	1 Cijfer Hoger Onderwijs	National level indicators list for higher education
ABF research	ABF research	Private company providing data services
Altmetric	Altmetric	Metrics and data about research, from Taylor & Francis
ARWU	Academic Ranking of World Universities	One of the widely used university rankings
Basismonitor arbeid en org	Basismonitor arbeid en organisatie	Database from SoFOkles about werk and organisations
CBS	Centraal Bureau van Statistiek	National agency for statistics (public data colletion)
cRIS	Current Research Information System	Information system used by the universities for research registration
CROHO	het landelijke centraal register opleidingen hog	National register for higher education teaching
CWTS	Centre for Science and Technology Studies	Agency from Leiden Unviersity for bibliometry and Scientometry
DAIR	Dutch Association for Institutional Research	Association for people from the universities regarding institutional research
DANS	Data Archiving and Networked Services	Data related institute under Royal Dutch Academy of Sciences
DataverseNL	DataverseNL	Online storage, sharing and registration of research data
DUO	Dienst Uitvoering Onderwijs	Education Executive service of Dutch government
Eurostat	Statistical office of the European Union	Statistical information on european level
IvhO	Inspectie van het Onderwijs	National agency for the inspection of education in the Netherlands
KUOZ	Kengetallen Universitair Onderzoek	Dataset for figures about university research
Loonaangifteketen	Loonaangifteketen	Government agency for public data (especially about income)
NAE	Nationaal Alumni Enquete	National level Alumni Survey
NARCIS	National Academic Research and Collaborations	National portal for information about researchers and their work
NRO	Nationaal Regieorgaan Onderwijsonderzoek	Netherlands Initiative for Education Research
NSE	National Student Enquete	National Level Student Survey
NUFFIC	Nederlandse Organisatie voor internationaliseri	Dutch organisation for internationalisation of education in the Netherlands
NVAO	Nederlands-Vlaamse Accreditatieorganisatie (N	Accreditation Organisation of the Netherlands and Flanders
NWO	Nederlandse Organisatie voor Wetenschappelij	Netherlands Organisation for scientific research
OCW	Ministerie van Onderwijs, Cultuur en Wetensch	Dutch ministry of education, culture, and sciences
OECD	Organisation for Economic Co-operation and De	Statistical information about economics on global level
OSIRIS	Online student information system	Student information system used by the universities in the Netherlands
PURE	PURE	Research information system from Elsevier
PwC	PricewaterhouseCoopers	International advising company
Rathenau	Rathenau Instituut	Institute for Research and dialogue relating to the societal aspects of science, innovation and technology
SciVal	SciVal	Research performance visualisation system from Elsevier
SK123	Studiekeuze 123	Agency for helping students to make education choices
SoFokles	Sociaal Fonds voor de Kennissector	Social funds for academic jobmarket projects
Statline	Statistical dataset	Statistical databank from CBS
THE	Times Higher Education ranking	One of the widely used university rankings
VH	Vereniging van Hogescholen	Association of the applied universities in the Netherlands
VSNU	Vereniging van Nederlandse Universiteiten	Association of the scientific universities in the Netherlands
WOPI	Wetenschappelijk Onderwijs Personeel Informa	Scientific Teaching human resources register

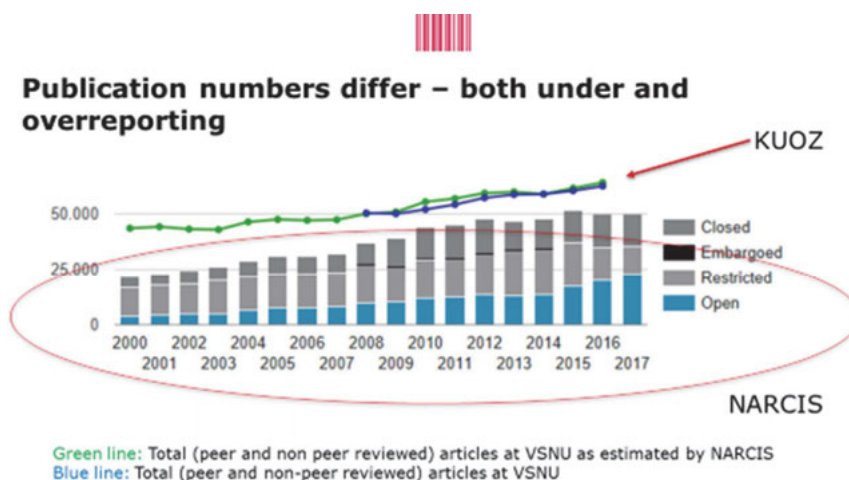*Table 1*. List of agencies with abbreviations and their translation or description.

Knowl. Org. 46(2019)No.5
Brief Communication

383

Figure 2. Difference in publication numbers in NARCIS and KUOZ.

NARCIS	KUOZ
Books, dissertations, conference proceedings	Scientific publications
Scientific output	
Articles	Scientific articles
	Not-peer reviewed articles
	Field articles (in Dutch:Vak)

Table 2. Difference in classification definitions and aggregation level used in NARCIS and KUOZ.

1.3 Alternatives to current RISs

Two options to solving the problems of fragmented governance of research data, with pro and counterarguments, are described below.

Option 1: Retrieve the data directly from the research information systems (CRISs) of universities to the VSNU via APIs.

Pro arguments:
– Easier for the universities to deliver the data;
– The work has already been partially done within RISs;
– Possible Machine2machine automation.

Counter arguments:
– The universities could lose control over the delivering the data;
– The mapping of CRISs with KUOZ (Standard Evaluation Protocol) SEP module must be arranged by the universities themselves;
– Not all universities use the same CRIS. A link among various RISs must be developed for each system.

Option 2: Request aggregated research output via NARCIS.

Pro arguments:
– Possible Machine2machine automation;
– Information on key research outputs can be retrieved at various levels (university, professor) continuously instead of periodically;
– A national (NARCIS) standard.

Counter arguments:
– RIS from the universities has API linking problems with NARCIS;
– Difference in definitions for research output between KUOZ and NARCIS;
– The universities must provide the extra data (Sector, type of publication) to NARCIS.

1.4 Need for a national level research data governance framework

A common data governance across the Netherlands is necessary to enable Dutch universities, research organizations and industry to benchmark education, research and innovation. It can be used to investigate the links between edu-

cation strategies, research practices and business outcomes. Furthermore, such framework could help researchers, companies and also the government with decision making in higher education. The definitive product of national level research data governance framework could be a national depo for the data collection, automation and as the "open data portal" for research output. Or as an alternative, it can be a federative model of data collection, where data are stored in various places, but there are common definitions and agreements on exchange of aggregated data. The current fragmented research output monitoring system can be replaced by such a common system.

A number of criteria have been considered for a common framework: it must be usable at the national level, it must be public, and it must encapsulate strategic goals such as the climate agreement and sustainable developments. In such a framework, all data can be mapped at a glance, arranged per theme. Considerable attention will be given to social impact in relation to research. It can be used as a good tool for participating universities to learn from each other's research data. The framework should have the possibility to incorporate the updated synchronization of data, new functionalities and topics. The information is free of charge and up-to-date and can, therefore, be used by anyone interested in the higher education sector: from journalist to student, from civil servant to professor.

1.4.1 Examples of a national level governance framework for research classification standards

Some countries are either developing or already have developed a national level governance framework to develop research classifications standards; for example, Canada and the Flanders region of Belgium (Legendre 2019; Vancauwenbergh and Poelmans 2019). The new Canadi-

an RIS is modeled on the established Australian and New Zealand Standard Research Classification (ANZSRC). It will align with international standards for collecting and reporting data on research (OECD's *Frascati Manual*). It will contain three dimensions of research, namely, type of activity, field of research, socio-economic objectives. Similarly, the Flemish research classification is based on classifying persons, organizations and projects. It contains research reporting, multiple discipline code lists, etc.

2.0 Suggestions from a brainstorming session with various stakeholders

A brainstorming session was held on the 15th of April 2019 among policymakers, data scientists and officials from various universities, institutes and government offices to discuss the issue of data in terms of findability, accessibility and interoperability. The session is summarized in tables 3-5 below.

3.0 Call for a working group to develop a national governance framework

We call for a working group to develop a common governance framework for research data in the Netherlands. The tasks of the working group could be to:

– develop consensus to collect, edit and use research data in the wake of GDPR;
– cluster publication to support research domains;
– determine the role of private parties in the research data exchange;
– develop key figures for policy information about research; and to
– focus on research community proposed public needs.

2.1 Findability: should we have one place for all data for higher education in the Netherlands?	Answer: No, federated model instead of one location
How do we focus on findability?	Theme oriented (Higher education, Research, Impact), Target group-oriented (e.g. administrators, coordinators, study advisors).
Criteria for findability of data:	Accountability by institutions (e.g. NRO, DANS for open access), Standard databases through quality check, Standardization of metadata, Recording of sources (1CHO, PhD students survey, KUOZ, WOPI, HODEX, CBS micro-data).

Table 3. Findability.

2.2 Accessibility: shall we have mutual agreements for data exchange?	Answer: Not only over exchange, but also over definitions should be agreed upon.
Purpose of accessibility:	Easy data delivery from and to universities.
Criteria for accessibility of data:	Fixed Metadata, Clarity about GDPR related issues, Source comparability with regard to internal insights and external communication, Data use through more mutual agreements.

Table 4. Accessibility.

Knowl. Org. 46(2019)No.5
Brief Communication

385

2.3 Interchangeability: should we have one set of definitions?	Answer: Not one set of definitions but clear description of definitions
Purpose of interchangeability	Use data and definitions that can be interchangeable
Criteria for exchangeability of data and metadata:	Use a common source per focus area (education, research and impact), A prognostic attitude instead of a retrospective attitude, Compatible with social accountability, Measurement of impact for 6 year period, Clear description and less granularity, National level quality check.

Table 5. Interchangeability.

In order to achieve these goals, the working group might consider following governance framework.

3.1 Classification

Currently available commercial data classification and aggregation systems are mentioned Legendre (2019) and Vancauwenbergh and Poelmans (2019). These classification systems give a first indication of number of publications, information about researchers, etc., subdivided into themes. The following considerations must be taken into account while developing a classification strategy for common data governance:

– Complete list of data providers, collectors and analyzers;
– Finding stakeholder institutes, research centers (CBS, Rathenau, publishers);
– Finding people involved in the efforts to maintain/develop RISs;
– International initiatives, like the Common European Research Information Format (CERIF) and EuroCris;
– Comparison of definitions among various RISs;
– Classifying available RISs: type of data, search hits, corresponding licenses;
– Designing a RISs terminology and RISs sections;
– Research and Valorization: Publications, PhDs, events, rankings, finance;
– Impact: Economic benefits, patents, collaboration with industry;
– Key Performance Indicators per section of information system of data (education, research and impact) for benchmarking and analysis for the sector-wide performance of universities. The indicators could give an overview of how domains within sectors are performing across a broader range of universities. The indicators should give information about domains within sectors compare against their peers abroad and the higher education sector on average.

3.2 Technology

Currently, every university has its own research information system. The universities will have to accommodate such interoperability among different metadata standards and schemas, extra metadata. A standalone website is suitable for using such a content management system supported by worldpress.org/.com. The following considerations must be taken into account while developing a technical framework for common data governance:

– Checking the publications for digital object identifier (DOI);
– Initiative by university libraries for the indicators for research and knowledge transfer domain;
– Include more information on metadata;
– Extracting the data from the various sources (scraping information from webpages).

3.3 Legal status

A common governance framework based on the open source is easier to deal with the legal issues for digital rights as compared to RISs with commercial interests. With its creative commons license it will provide public access to data and to official publications of universities regarding education, research and valorization.

4.0 Concluding remarks

Figure 1 shows that the information landscape on data on RIS is very diverse, even under the existence of CRIS—that is, already set standards for RIS. The implementation of these CRIS standards is still very diverse, partly because different platforms are used (table 6) (propriety tools and open ones); because content is created in different ways (from individual user created content to more centrally curated content); and the categorization schemes also vary—so data are not harmonized.

It is evident from the stakeholders session described above that there is a need for more complete, up-to-date, even timely information, both for science policy on a national level and for its application at university level. The VSNU is taking on a central coordinating role in this process, which is a practical policy process and still needs to be informed by research questions such as those con-

386

Knowl. Org. 46(2019)No.5
Brief Communication

University short-name	CRIS	Link
TUD	PURE	https://pure.tudelft.nl/portal/
EUR	METIS	https://www.eur.nl/node/22383
RU	METIS	https://www.ru.nl/ubn/bibliotheek
TUE	PURE	pure.tue.nl
UM	PURE	https://library.maastrichtuniversity.nl/cris-support/
LEI	LUCRIS	https://lucris.weblog.leidenuniv.nl/
UvA	PURE	http://uba.uva.nl/diensten/voor-onderzoekers/pure/pure.html
VU	PURE	https://research.vu.nl/en/
OU	PURE	https://research.ou.nl/
RUG	PURE	https://www.rug.nl/library/support/pure/
WUR	PURE	https://www.wur.nl/en/Value-Creation-Cooperation.htm
TiU	PURE	https://research.tilburguniversity.edu/en/
UT	PURE	https://www.utwente.nl/nl/ris/

Table 6. List of current research information systems (CRIS) used at universities.

cerning the interoperability and functionality of classifications, hence our contribution in this journal.

Notes

1. De Vereniging van Universiteiten (VSNU): Association of Universities in the Netherlands.
2. Ministerie van Onderwijs, Cultuur en Wetenschap (OCW): Ministry of Education, Culture and Science.

References

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

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

Appendix

Facts and Figures pages from various universities

https://vsnu.nl/nl_NL/feiten-en-cijfers-universiteiten.html

https://www.tudelft.nl/en/about-tu-delft/facts-and-figures/

https://assets.tue.nl/fileadmin/content/universiteit/universiteit/feiten_cijfers/TUe.WEB.pdf

https://www.uu.nl/en/organisation/profile/facts-and-figures

https://www.uva.nl/en/about-the-uva/about-the-university/facts-and-figures/facts-and-figures.html?1561919410238

https://www.acta.nl/en/about-acta/organization/facts-figures/index.aspx#

https://fsw.vu.nl/en/about-the-faculty/facts-and-figures/index.aspx

https://sbe.vu.nl/nl/over-de-faculteit/missie-en-visie/facts-figures/index.aspx

https://www.ru.nl/english/about-us/our-university/facts-figures/

Various Data sources:

– National level Data sets: 1cijfer HO, WOPI, KUOZ, CROHO
– Survey data: National Student Survey, National Alumni Survey (NAE)
– Data collectors / users : OCW, VSNU, DUO, Nuffic, CBS, Eurostat, NVAO
– Ranking agencies: THE, ARWU, CWTS,
– Awarding/Funding agencies: NWO, Horizon 2020