Framework document Steering group VSNU/NFU/NWO – Elsevier

Name Pilot/Service: Data Monitoring support service

Short description of the pilot/service:

Data Monitor is about the F in FAIR data. Making sure data is findable and trackable. Through the Data Monitor service access is provided to the data via a reporting interface, an (optional) public web page as well as an API so that it can be integrated with Research information systems. With Data Monitor, universities can more easily showcase and reward researchers and projects that have gone through the effort of sharing their data openly, without adding an additional admin burden to the researchers.

This service provides better mapping as to the location of research data with a system that:

- Automatically inventories where datasets are located, provided the public data repository can be tracked. This helps institutions showcase and report on open data sharing, and check DMP's (data management plans)
- Tracks metadata of these research data sets. E.g. particularly the description, authors, and related research article
- Integrates with the partner's CRIS to automatically upload this metadata as part of the standard institutional CRIS curation process

With this pilot, we want to learn:

- How useful the service is for the faculty and staff (how representative is it)
- How useful the service is for RDM team (how much it helps to save time and engage with faculty)

This service will also enable participating institutions to evaluate how their various faculties and institutes go about their data management plans and protocols. The objective would be to increase visibility/awareness and promote the importance of FAIR data.

More details in Appendix 1.

1. (a) Participating institutions Participation in the Professional Services is at each Institution's sole discretion and a pilot shall only commence if there is a minimum participation by at least three Institutions *	Evaluation YES NO	Evidence and Comments
Are at least 3 institutions involved in the pilot?	YES	11 institutions have showed interest in this pilot, 9 were sent a draft SoW and 4 already signed it (more details in Appendix 2)
Evidence of how and when other institutions can join	YES	Any institution can join the pilot at any point in time.
1. (b) Interoperability and vendor neutrality Elsevier shall use all reasonable efforts to ensure that the Professional Services are interoperable, both on the input (uploaded) and output side (created) *	Evaluation YES NO	Evidence and Comments

Use of open identifier systems	YES	Datasets are identified in Data Monitor via their source persistent identifiers: this includes DOIs, handles, PURLs and others. Articles associated to datasets are mostly identified via DOIs. Data Monitor supports open taxonomies and URLs for metadata fields such as copyright and license. Data Monitor supports identifiers for dataset contributors and their affiliations as well, including ORCID, ROR as well as Scopus, SciVal and Mendeley identifiers. Data Monitor enrichment pipeline is designed to maximize author and institutional attribution, as well as enabling linking with entities in Current Research Information Systems such as Pure. Data Monitor uses standard Article and Dataset's DOIs To identify an institution, Data Monitor uses a broad range of identifiers (incl. ROR) which are provided in the source metadata. During the enrichment process, at the moment the SciVal and Scopus institutional identifiers are used. We are looking to extend this over time towards also using the ROR identifier in the enrichment process, when it has reached sufficient maturity.
Use of standardised metadata schemas	N/A	Data Monitor aims to support the OpenAIRE metadata schema, regarded by many as the most accepted standard at the moment. To date, we already support the vast majority of the mandatory metadata fields, and we are working to support the entire set in 2021. Furthermore, we are tracking efforts in the community to expand our support of future standards, which are still in the making (see https://www.rd-alliance.org/groups/open-science-graphs-fair-data-ig)
Existence of a well-documented API and open data-dump function	YES	https://datasearch.elsevier.com/api/docs#/search/search_1
Ability to export data in a variety of formats	YES	Excel, Csv, JSON
Ability for other commercial parties to join	YES	Any other 3rd commercial party can join the pilot and integrate with the dedicated API (or another between parties agreed mechanism to interlink) acknowledging that the data will not be copied in full for its own specific purposes, and/or for purposes not contributing to the pilot. Furthermore, Data Monitor provides a dedicated API that allows 3rd party repositories to be indexed, thus extending Data Monitor's coverage of over 2000 repositories.

2. **Transparency, inclusion and collaboration** The Services and resulting Deliverables are aimed to make science and research more transparent, efficient, inclusive, openly and freely accessible, and collaborative. *	Evaluation YES NO	Evidence and Comments
Provenance on how and where metadata was derived	YES	Data Monitor indexes over 2,000 data repositories either directly or via DataCite or OpenAire. This translates to tens of millions of indexed datasets. Our complete list of indexed repositories can be provided upon request, or accessed via our API at: https://datasearch.elsevier.com/api/docs
Descriptions of workflows that result in indicators, metrics and/or other relevant outcomes will be open and transparent. These will demonstrate, for example calculation steps, search strings used to define entities, etc.	NA	-
Description of the services used to create metadata	YES	More details in Appendix 1 (pipeline slide)
Insights and lessons published with Open Access license	YES	This is at the discretion of the partners. We would love to be a co-auther to the publishing of these insights and learnings.
Will the pilot contribute to open science?	YES	It will help institutions to track, report and showcase better on Open Data (2nd pillar of Open Science) whilst avoiding an extra admin burden on researchers.
Demonstration of connection to non-Elsevier products	YES	Data Monitor already connects to thousands of non-Elsevier repositories. On top of these thousands non-Elsevier repositories, Data Monitor also connects to Elsevier's Mendeley Data repository
3. **Access to research data and metadata** Elsevier will give enduring access during the Term to all (research) data, including metadata, analytics and information *	Evaluation YES NO	Evidence and Comments
Describe the ownership / licencing of data made as part of the service	YES	The specific licence of the datasets depends on the metadata provider (data repositories), but all of it can be in the public domain (all of it is OA) The Metadata links to public data sources and there is no transfer of data ownership. Once DataMonitor metadata

		has been provided, it will be put into the system of record of the institution. The institution then decides what do with this data, so if it is open, it remains open. This way the data is also available for the NL OKB under development, if this is needed.
Describe how access (institutional and / or public) to the data will be set-up during the term; this section will also indicate cases where certain data is not publicly accessible.	YES	Data Monitor provides access to the data via a reporting interface, a public web page as well as an API and it is natively integrated with Pure and can be integrated with other Research information systems.
4. Data portability Institution shall be entitled to transfer the data provided, uploaded or created to its own or to a third party host environment *	Evaluation YES NO	Evidence Comments
Evidence on how data transfer is possible	YES	The data is portable and/or remains in the institutional information system when the service is cancelled/discontinued.
How can an institution withdraw data?	NA	NA
5. Intellectual property *	Evaluation YES NO	Evidence Comments
Details on IP related to data provided, created or enriched	NA	NA
6. Additional considerations	Evaluation YES NO	Evidence Comments
What processes will be put in place to evaluate the service during and at the end of pilot	YES	It is known that tracking open data is a hard problem, both in terms of the completeness and quality of the data. We have designed a quantitative evaluation framework that aims at measuring both aspects and track progress over time. Therefore, the pilots are being evaluated by the institutions on both these criteria, without knowing up-front how accurate or precise they expect the feed to be (there is currently no benchmark)
Terms of use of the deliverables during and after contract period	YES	The deliverables of this pilot are stated in the SoW. If participating institutions so wish, service is available as part of the agreement until December 31, 2024. In case participating institutions do not want to continue to use the service, all public data serviced by Data Monitor will remain on that website (if participating institutions so wish)

Pilot project team	YES	Names of the partners' contacts detailed in the respective SoWs

* For the full text, please refer to the contract.

Approved by the VSNU/NFU/NWO-Elsevier steering group	Date: 2 March 2021