Framework document steering group UNL-NFU-NWO-Elsevier

Name Pilot/Service: Rare Disease Analytics

Short description of the project:

Introduction

Rare diseases are, despite its name, very frequent: about 7000 rare diseases hit approximately 30 million people in Europe alone (1.5 million in the Netherlands). Though there is quite some research done into many rare diseases, their findings are hidden in the enormous number of medical publications. Their metadata is usually not adequately indexed by search engines like PubMed or Embase and therefore difficult to discover.
The NFU assigned 300 expertise centers for these rare diseases, but their research and outcome (articles) overview are incomplete. By refining the indexing in Scopus based on an Erasmus prototype the metadata of these articles will be improved and linked to the relevant centers.

The aim of this pilot is to provide open access to insights (articles, metadata, other information) about rare (or orphan) disease research activities in the Netherlands to researchers, institutions, patients, funders and other stakeholders in health care, stimulating further research (patient data) to rare diseases and contributing to public knowledge.

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	The project is being executed with help and support of all NFU members.
Evidence of how and when other institutions can join	YES	2nd and 3rd tier hospitals could join as well, as well as independent research institutions in medicine.
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	Orpha (see orpha.net) and DOI

Use of standardized metadata schemas	N/A	
Existence of a well-documented API and open data-dump function	N/A	If an API is required, it will be one using standard technology
Ability to export data in a variety of formats	YES	Most likely all that is needed will be lists connecting Orpha codes with DOIs.
Ability for other commercial parties to join	YES	Members can share their Orpha codes per DOI with anyone.
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	The current plan is to use a proprietary NLP-based algorithm (of which the principle will be explained in exact wording) to match Orpha codes and DOIs.
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.	YES	The dashboard will be created with a representation of the community of rare disease experts. (about 50% of academic researchers in UMC's are likely involved with rare disease research.) They test data and provide desired specs for the dashboard
Description of the services used to create (meta)data		NLP-based algorithm (of which the principle will be described/published in relevant conference proceedings) creates search queries for each Orpha code. These queries will run over Scopus data and generate lists of DOIs per Orpha code.
Insights and lessons published with Open Access license	NA	
Will the pilot contribute to Open Science?	YES	The Creation of an open system that provide insights in Rare Disease research progress will create

		- efficiencies for the evaluation processes already in place
		- visibility of outputs and outcomes per disease or disease cluster for:
		  o research groups
		  o institutions
		  o funders
		  o Governments
		  o Patient Groups
		- visibility of research efforts and experts for patients
		- visibility of developments, trends and experts for:
		  o pharmaceutical companies
		  o health insurers
		- This project will support Evidence Based Medicine, (medical practice or care that emphasizes the practical application of the findings of the best available current research)

Demonstration of connection to non-Elsevier products	YES	DOIs are used by all vendors in research eco system. The Orpha codes are publicly accessible. If an API is required, we will provide one that uses standard technology.
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 Comments
Describe the ownership / licensing of data made as part of the service	YES	The connection between Orpha codes and Dutch DOIs will be made available on a public site and to the individual members.
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 access.	NA	

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	n/a	Lists of Orpha-DOI combinations will be provided to the institutions; format will be agreed by the parties.
How can an institution withdraw data?	N/A	
5. Intellectual property	Evaluation YES NO	Evidence Comments
Details on IP related to data provided, created, or enriched		The service will use a proprietary NLP-based algorithm to match Orpha codes with DOIs. The principles of the algorithm will be described/published in two conference papers and an upcoming journal article.
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		The service will be a co-creation with NFU and will be supervised by senior NFU officials. In alignment with the participating institutions, a review process will be setup to test the outcomes of the service.
Terms of use of the deliverables during and after contract period		The deliverables of this pilot are stated in the SoW. If participating institutions so wish, the service is available until the end of 2024. All data provided to the institutions remains with the institutions also after the end of this agreement.
Pilot project team		The service will be a co-creation with NFU and will be supervised by senior NFU officials