FOSB WG Metadata en Standaardisatie

*Sense of urgency*

Anno 2020 is het Vlaamse onderzoekslandschap vrij divers inzake de identiteit en de kenmerken van de onderzoek uitvoerende organisaties (RPO's), alsook de wijze van organisatie en maturiteit van elke RPO met betrekking tot het beheer van onderzoeksgegevens. Deze diversiteit en het verschil in maturiteit geldt ook binnen en tussen de verschillende onderzoeksdisciplines waarin de Vlaamse RPO's actief zijn. Momenteel worden alle Vlaamse RPO's door middel van decreten en overeenkomsten gevraagd om metadata informatie te verstrekken aan het FRIS-onderzoeksportaal (researchportal.be). Dit wordt formeel uitgedrukt binnen het BOF- en IOF-decreet (van toepassing op de Vlaamse Universiteiten), overeenkomsten (DOSP voor de hogescholen, convenanten voor de strategische onderzoekscentra, ...).

Het FRIS-portaal ontsluit momenteel informatie over onderzoekers, onderzoeksgroepen, publiek gefinancierde projecten en alle publicaties vanaf 2008 en is toegankelijk voor iedereen. Indien het FRIS-portaal ook de metadata informatie over onderzoeksgegevens wil ontsluiten op een eenduidige manier, is het van het grootste belang dat alle RPO’s gebruik maken van eenzelfde metadatavelden die semantisch op eenzelfde manier ingevuld worden door alle stakeholders. Hierdoor wordt een perfect begrip van de informatie op het FRIS-portaal mogelijk, wat zorgt voor een hoger hergebruik van de onderzoeksgegevens, wat de return on investment voor het creëren van de onderzoeksgegevens verhoogd. Op dit moment hebben sommige RPO's hun metadatamodellen al in gebruik, maar veel RPO's zijn op dit moment nog op zoek welk metadatamodel en welke disciplinaire standaarden bij voorkeur gebruikt worden. Het is dan ook het momentum om een gemeenschappelijk metadatamodel te bepalen dat semantisch op eenzelfde manier wordt ingevuld door alle Vlaamse stakeholders.

*Strategische doelstellingen*

De FOSB WG Metadata en Standaardisatie richt zich daarom op de ontwikkeling van een semantisch beschreven generiek metadatamodel voor onderzoeksgegevens, dat gebaseerd is op bestaande metadatamodellen (bv. OpenAire, DataCite, ...) en rekening houdt met de aanbevelingen van de RDA. Tegelijkertijd worden de metadatavelden die de onderzoeksgegevens beschrijven, getoetst aan de FAIR-gegevensprincipes (Force11), zodat de metadata vindbaar, toegankelijk, interoperabel en herbruikbaar zijn. Het resulterende metadatamodel zal door de Vlaamse onderzoeks- en financieringsorganisaties gebruikt worden om de metadata op een semantisch geharmoniseerde manier aan FRIS aan te bieden, zodat de metadata op een gestandaardiseerde manier door alle informatieverstrekkers op een eenduidige manier kunnen worden ontsloten. Aangezien de metadata op het FRIS-portaal toegankelijk zijn voor iedereen via een Open Data-licentie, zal het gebruik van eenzelfde taal voor het definiëren van onderzoeksmetadata ertoe leiden dat geautomatiseerde gegevensextractie mogelijk wordt en de metadata machine-readable worden, wat het hergebruik van onderzoeksgegevens tussen alle Vlaamse stakeholders maximaliseert en meer analyse mogelijkheden toelaat via gebruik van data science methoden. Om de implementatie te ondersteunen, zal de FOSB WG Metadata en Standaardisatie semantisch beschreven concordantietabellen ter beschikking stellen wanneer er afwijkingen bestaan tussen het generieke metadatamodel en het model van de informatieverstrekkende instellingen.

Daarnaast zal de FOSB WG Metadata en Standaardisatie een gelijkaardige aanpak aanwenden voor disciplinaire metadata voor wat betreft de ontwikkeling van een semantisch beschreven metadatamodel en de ondersteuning van de implementatie. De uitdaging hier is om een semantisch beschreven standaard voor disciplinaire metadata af te bakenen in de grote verscheidenheid aan disciplinaire standaarden die vandaag de dag bestaan bij alle betrokken (disciplinaire) stakeholders, en dit in lijn met de ontwikkelingen binnen EOSC. Hiertoe zal een gelijkaardige methodologie gehanteerd worden zoals hierboven beschreven om te komen tot disciplinaire standaarden. Ten slotte zal de WG ook de ontwikkelingen op Europees en internationaal niveau op het gebied van metadata standaarden volgen, met nadruk op de ontwikkelingen binnen EOSC vermits FRIS zal fungeren als connectie tussen de metadata repositories van de RPO’s en EOSC.

*Krachtlijnen*

1. Ontwikkeling en ondersteuning van de implementatie van een semantisch beschreven generiek metadatamodel voor metagegevens over onderzoeksgegevens

2. Ontwikkeling en ondersteuning van de implementatie van een semantisch beschreven disciplinair metadatamodel voor onderzoeksgegevens.

3. Toezicht op en aanpassing aan Europese en internationale trends, met name op het gebied van het EOSC

*Roadmap*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Doelstellingen – timing (*ongoing* – *done*) | 2020 | | | 2021 | | 2022 | 2023- ... |
|  | Q2 | Q3 | Q4 | Q1/Q2 | Q3/Q4 |  |  |
| 1. Ontwikkeling en ondersteuning van de implementatie van een semantisch beschreven **generiek** metadatamodel voor metagegevens over onderzoeksgegevens | | | | | | | |
| * 1. *Ontwikkeling van een generiek metadatamodel voor de beschrijving van onderzoeksgegevens* | | | | | | | |
| 1.1.1 Eerste ontwerp van een model op basis van gemeenschappelijke standaarden |  |  |  |  |  |  |  |
| 1.1.2 Review en verfijning van metadatavelden en semantiek |  |  |  |  |  |  |  |
| 1.1.3 Evaluatie van metadata velden mbt de FAIR-principes |  |  |  |  |  |  |  |
| 1.1.4 Validatie van het generieke metadatamodel |  |  |  |  |  |  |  |
| * 1. *Ondersteuning van de implementatie van het generieke metadatamodel voor de beschrijving van onderzoeksgegevens* | | | | | | | |
| 1.2.1 Identificatie van discrepanties tussen het generieke metadatamodel en de bestaande metadatamodellen van de informatieverstrekkende instellingen |  |  |  |  |  |  |  |
| 1.2.2 Opstellen van semantisch beschreven concordantietabellen voor informatieverstrekkers |  |  |  |  |  |  |  |
| 1.2.3 Ondersteuning van de implementatie van het generieke metadatamodel in FRIS en in alle Vlaamse informatieverstrekkende instellingen (in overeenstemming met het tempo en de maturiteit van elke instelling met betrekking tot RDM) |  |  |  |  |  |  |  |
| 1.2.4 Monitoring en opvolging van het generieke metadatamodel |  |  |  |  |  |  |  |
| 1. Ontwikkeling en ondersteuning van de implementatie van semantisch beschreven **disciplinaire metadata** over onderzoeksgegevens | | | | | | | |
| * 1. *Ontwikkeling van disciplinaire normen voor de beschrijving van onderzoeksgegevens* | | | | | | | |
| 2.1.1 Inventarisatie van disciplinaire metadata standaarden |  |  |  |  |  |  |  |
| 2.1.2 Per discipline:  - afbakenen van de disciplinaire standaard  - beoordeling & verfijning  - toevoeging van semantische beschrijvingen  - toetsing aan de FAIR-principes  - validatie van de disciplinaire normen |  |  |  |  |  |  |  |
| *2.2 Ondersteuning van de toepassing van de disciplinaire normen voor de beschrijving van onderzoeksgegevens* | | | | | | | |
| 2.2.1 Beoordeling van de integratie van discipline-specifieke standaarden in het generieke FRIS-metatamodel |  |  |  |  |  |  |  |
| 2.2.2 Identificatie van discrepanties tussen de disciplinaire normen en de instellingsspecifieke normen, met inbegrip van het opstellen van semantisch beschreven concordantietabellen voor informatieverstrekkers |  |  |  |  |  |  |  |
| 2.2.3 Het ondersteunen van de implementatie van de disciplinaire normen in FRIS en in alle Vlaamse informatieverstrekkende instellingen (in overeenstemming met het tempo en de maturiteit van elke instelling met betrekking tot RDM) |  |  |  |  |  |  |  |
| 2.2.4 Toezicht op en follow-up van de disciplinaire normen |  |  |  |  |  |  |  |
| 1. **Monitoring en aanpassing** aan Europese en internationale trends, met name op het gebied van het **EOSC** | | | | | | | |
| 3.1 Actieve monitoring van Europese en internationale trends, met name op het gebied van het EOSC |  |  |  |  |  |  |  |
| 3.2 Aanpassing en verfijning van het generieke metadatamodel en de disciplinaire richtlijnen wanneer dat nodig is |  |  |  |  |  |  |  |