

Revisão sistemática da literatura sobre o envolvimento de investigadores em estudos GDI

João Aguiar Castro, Yulia Karimova, Cristina Ribeiro
Faculdade de Engenharia da Universidade do Porto / INESC TEC

Motivação

(1) Ao **envolver os investigadores** na seleção de conceitos para descrever os dados que produzem, os investigadores vão criar **metadados de maior qualidade**.

(2) Ao retirarem **benefícios** da descrição de dados (**reutilização**), os investigadores **vão estar mais motivados a participar em actividades GDI**.

Metadados FAIR

(R) reutilização: dependente de metadados específicos ao domínio de produção

Limitações

- Modelo de publicação tradicional
- Falta de profissionais especializados
- Investigadores pouco familiarizados com GDI

Qual o envolvimento dos investigadores em estudos GDI relacionados com metadados?

Expressão pesquisa

Scopus®

(1) **Metadata** no título ou keyword

(2) **{research data}** ou **{scientific data}** no título, abstract ou keyword

(3) **metadata** {standard} ou {schem*} ou **ontology** ou **vocabulary**

(TITLE (metadata) OR KEY (metadata) AND TITLE-ABS-KEY ({research data} OR {scientific data}) AND ALL ({metadata standard} OR {metadata schema} OR {metadata scheme} OR ontology OR vocabulary))



Set feed

132 Referências

11 positivos
(8,3%)

```
@CONFERENCE{Hoekstra201328,  
author={Hoekstra, R. and Groth, P.},  
title={Linkitup: Link discovery for research data},  
journal={AAAI Fall Symposium - Technical Report},  
year={2013},  
note={cited By 2},  
url={https://www.scopus.com/inward/record.uri?eid=2-  
s2.0-84898926852&partnerID=40&md5=f51a578e61635cac24f050ae055230a0},  
abstract={Linkitup is a Web-based dashboard for enrichment of research output published via industry grade data repository  
services. It takes metadata entered through Figshare.com and tries to find equivalent terms, categories, persons or entities on the  
Linked Data cloud and several Web 2.0 services. It extracts references from publications, and tries to find the corresponding Digital  
Object Identifier (DOI). Linkitup feeds the enriched metadata back as links to the original article in the repository, but also builds a  
RDF representation of the metadata that can be downloaded separately, or published as research output in its own right. In this  
paper, we compare Linkitup to the standard workflow of publishing linked data, and show that it significantly lowers the threshold  
for publishing linked research data. Copyright © 2013, Association for the Advancement of Artificial Intelligence. All rights  
reserved.},  
document_type={Conference Paper},  
source={Scopus},  
}
```

```
@CONFERENCE{Wira-Alam201256,  
author={Wira-Alam, A. and Dimitrov, D. and Zenk-Möltgen, W.},  
title={Extending basic dublin core elements for an open research data archive},  
journal={Proceedings of the International Conference on Dublin Core and Metadata Applications},  
year={2012},  
note={cited By 2},  
url={https://www.scopus.com/inward/record.uri?eid=2-  
s2.0-84879681987&partnerID=40&md5=5c0ae06a397af702d4f30c16a69c861f},  
abstract={In our project "DATORIJUM", we intend to provide a simple, open research data repository which focuses on social  
science research data. We encourage researchers to deposit their data and disseminate them among communities or academic  
partners. One of the key problems for longterm archiving is ensuring that the metadata elements are consistent and compatible  
with other standards. This paper discusses the use of basic Dublin Core elements with some simple extensions for structuring the  
data at study level. Moreover, we also depict the interplay between the emerging combination and the DDI metadata elements,  
particularly DDI-Lifecycle, and the possibility of using RDF to bring the data into the Linked Open Data Cloud.},  
author_keywords={Data documentation initiative (DDI); DATORIJUM; Dublin core; Linked open data; Metadata; Social science  
research data},  
document_type={Conference Paper},  
source={Scopus},  
}
```

Artigos

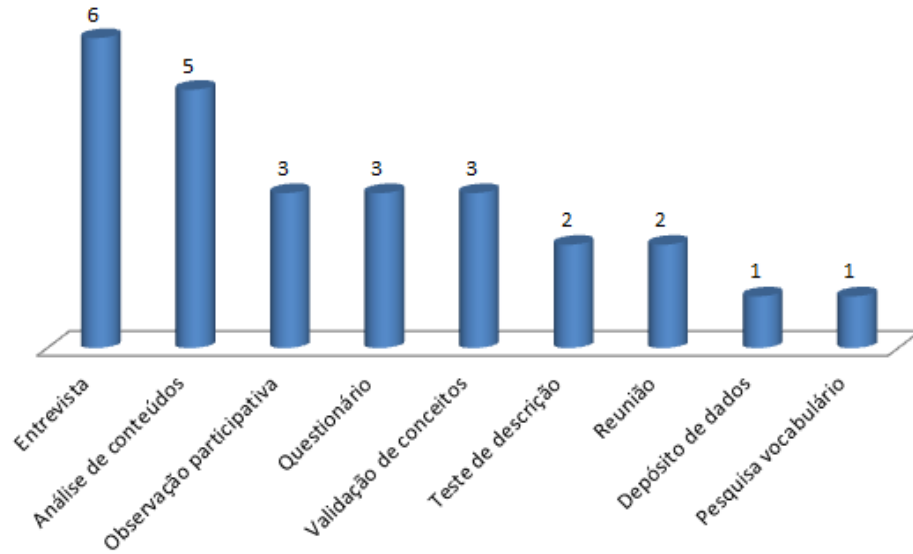
3	SPECTRA: The Deposition and Validation of Primary Chemistry Data in Digital Repositories	3_SPECTRA
31	How institutional factors influence the creation of scientific metadata	31_Institutional_Influence
32	Considering personal organization: Metadata practices of scientists	32_Personal_Organization
51	Research Data and Metadata Curation as Institutional Issues	51_Institutional_Issues
59	LabTablet: Semantic Metadata Collection on a Multi-domain Laboratory Notebook	59_Labtablet
60	Beyond Inspire: An ontology for biodiversity metadata records	60_Inspire_Ontology
75	BioWes - from design of experiment, through protocol to repository, control, standardization and back training	75_BioWes
89	Engaging Researchers in Data Management with Labtablet, an Electronic Laboratory Notebook	89_Labtablet_ELN
90	Ontologies for research data description: a design process applied to Vehicle Simulation	90_Design_Process
98	Descriptive Metadata for Scientific Data Repositories: A comparison of Information Scientists and Scientist Organizing Behaviors	98_Organizing_Behaviors
106	Linked Data for the Natural Sciences: Two Use Cases in Chemistry and Biology	106_Linked_Data

Motivação e desafios

3_Spectra	Aferir práticas dos investigadores no arquivo e disseminação dos dados
31_Institutional_Influence	De que forma surgem fricções na criação de metadados
32_Personal_Organization	Como os investigadores gerem os dados e o impacto na comunidade da ciência da informação
51_Institutional_Issues	Avaliação de suporte institucional à gestão de dados e metadados ao nível disciplinar e institucional
59_Labtablet	Vantagens do uso de ferramentas de complemento aos cadernos de laboratório
75_BioWes	Descrição de métodos experimentais geralmente está incompleta.
90_Design_Process	Ausência de modelos de metadados em domínios específicos.
98_Organizing_Behaviors	Como especialistas da informação e investigadores organizam os dados
106_Linked_Data	Barreiras à publicação de dados como LOD são altas

	Participants	Domains
3_Spectra	-	Química
31_Institutional_Influence	3 instituições	Geociências Biodiversidade Física
32_Personal_Organization	7 investigadores	Biologia Evolucionária
51_Institutional_Issues	14 investigadores CENS 19 investigadores + staff UCAR	Geociências Biodiversidade
59_Labtablet	2 investigadores	Química Biologia Evolucionária
60_Inspire_Ontology	Grupo de investigação	Biodiversidade
75_BioWes	Infraestrutura internacional	Biomedicina
89_Labtablet_ELN	1 investigador	Ciências Sociais
90_Design_Process	1 investigador	Simulação de Veículos
98_Organizing_Behaviors	16 investigadores	Biologia
106_Linked_Data	Dois investigadores	Química Biologia

Técnicas utilizadas



51_Institutional_Issues	Entrevista Análise conteúdo Observação participativa
106_Linked_Data	Reunião Entrevista Análise Conteúdo Pesquisa de Vocabulário

Práticas dos investigadores

3_Spectra	Desconhecimento de repositórios disponíveis.
31_Institutional_Influence	Os processos são ad hoc e irregulares (CENS e ESG), permanentes no LTER.
32_Personal_Organization	Investigadores usam, sobretudo, metadados e esquemas de organização pessoais.
60_Inspire_Ontology	O grupo usa um geo-portal para partilhar metadados do projeto.
75_BioWes	Terminologia própria dos grupos geralmente utilizada.
89_Labtablet_ELN	Conhecimento de recomendações mas sem as colocar em prática.
90_Design_Process	Variáveis da simulação em documento de texto.
98_Organizing_Behaviors	Investigadores produzem metadados mais especializados que especialistas da informação.

Recomendações e sugestões

3_Spectra	Metadados específicos ao contexto são acrescentados para permitir reutilização.
31_Institutional_Influence	O grau de institucionalização dos metadados influencia as práticas e necessidades GDI.
32_Personal_Organization	Padronização na forma como biólogos organizam os dados.
59_Labtablet	Caderno de laboratório eletrônico é útil para capturar dados de oportunidade.
89_Labtablet_ELN	A metodologia é tida como um item essencial na descrição.
90_Design_Process	Modelos de metadados devem resultar da colaboração entre curadores e investigadores.
98_Organizing_Behaviors	Será útil considerar esquemas de metadados fáceis de adaptar por investigadores e especialistas da informação.
106_Linked_Data	A colaboração próxima entre investigadores e curadores é muito importante para o sucesso dos projetos.

Conclusões gerais

- A produção de metadados FAIR é limitada;
- Investigadores são quem melhor pode descrever os dados que produzem;
- Necessidade de envolver investigadores e curadores em iniciativas GDI.