



Uma Representação Semântica para Planos de Gestão de Dados

João Cardoso^a, Leyla J. Garcia^b, Fajar J. Ekaputra^c, Marie C. Jacquemot^d, Marek Suchánek^e, Toasz Miksa^{c,f}, José Borbinha^a

^aINESC-ID, [joao.m.f.cardoso, jlb]@tecnico.ulisboa.pt

^bZB Med Information Centre for Life Sciences, lj.garcia.co@gmail.com

^cTU Wien, fajar@juang.id

^dDMP OPIDoR, Inist-CNRS, marie-christine.jacquemot@inist.fr

^eFaculty of Information Technology, Czech Technical University in Prague, marek.suchanek@ds-wizard.org

^fSBA Research, tmiksa@sba-research.org

Resumo

Gerir de forma sistemática grandes volumes de dados de investigação é um dos principais desafios que afectam actualmente a comunidade científica. Os Planos de Gestão de Dados (PGD) surgem como uma ferramenta que permite à comunidade atacar este desafio. Um PGD é um documento que tem como objectivo documentar as técnicas, métodos e políticas de como os dados são geridos durante o seu ciclo de vida. Em adição também descreve as associações entre actividades de gestão de dados e as entidades ou pessoas responsáveis pela sua execução.

O conceito de PGD Accionável por Máquinas (PGDma) é uma evolução do conceito original de PGD, que procura mitigar algumas das limitações dos PGD. O PGDma têm como principal propósito a representação de um PGD de forma a que o consumo da informação não seja exclusiva para humanos, mas que também seja uma possibilidade para sistemas automatizados. Esta representação accionável por máquinas abre um gama de possibilidades de aplicação, tais como: (1) A utilização dos PGDma como um meio de troca de informação entre sistemas; (2) Uma maior integração dos PGD em fluxos de gestão de dados existentes; e (3) A automação de actividades de gestão de dados.

Esta flash talk apresenta um trabalho desenvolvido no contexto do DMP Common Standards (DCS) working group da Research Data Alliance (RDA), em que se procuraram definir um standard para a definição de conceito de PGD e formatos de representação de PGDma. Em particular foca-se na descrição da DMP Common Standard Ontology (DCSO) que permite a

representação de PGDma através do uso de tecnologia semântica.

Palavras-chave: Planos de Gestão de Dados, Ontologia, Automação

Designação do projeto/infraestrutura/iniciativa

DMP Common Standard Ontology (DCSO)

Público-alvo

O público-alvo desta iniciativa engloba todos os interessados (e.g. cientistas, investigadores, gestores de projeto, etc.) no tópico dos PGD. Em particular acerca dos métodos de representação de PGDma e no seu potencial uso futuro.

Ligações web úteis

<https://www.rd-alliance.org/>

<https://www.rd-alliance.org/groups/dmp-common-standards-wg>

<https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/tree/master/ontologies>