



Validador RCAAP: ferramenta para promover a interoperabilidade e a qualidade dos metadados

José Carvalho^a, Raquel Truta^b, Pedro Príncipe^c

^aUniversidade do Minho, Portugal, jcarvalho@sdum.uminho.pt

^bUniversidade do Minho, Portugal, raquel@sdum.uminho.pt

^cUniversidade do Minho, Portugal, pedroprincipe@sdum.uminho.pt

Resumo

O Validador RCAAP é uma ferramenta desenvolvida no âmbito do projeto RCAAP – Repositório Científico de Acesso Aberto de Portugal com o objetivo de aferir o grau de conformidade de um repositório ou revista com as diretrizes definidas no campo de ação do projeto. Apesar da aplicação de normas padronizadas, como o Dublin Coreⁱ e o OAI-PMHⁱⁱ, as representações locais comprometem, muitas vezes, a qualidade dos metadados e a interoperabilidade técnica. No âmbito do projeto RCAAP utilizam-se as diretrizes DRIVER como o padrão de interoperabilidade dos repositórios portugueses. Para além de assegurarem a necessária uniformização entre os repositórios e revistas integrantes do RCAAP, garantem e facilitam a integração e a interoperabilidade. Esta apresentação pretende expor as características e objetivos do Validador RCAAP e destacar os desenvolvimentos da segunda versão desta ferramenta, que oferece aos gestores de repositórios e revistas um serviço capaz de promover a qualidade e harmonia dos metadados, permitindo a interoperabilidade à escala nacional ou internacional.

Palavras-chave: repositório, revista, qualidade, diretrizes e interoperabilidade.

Conteúdo da apresentação

O projeto RCAAP – Repositório Científico de Acesso Aberto de Portugalⁱⁱⁱ – teve início em 2008 com a disponibilização de um portal de pesquisa (o portal RCAAP – www.rcaap.pt), para facilitar o acesso a conteúdos de acesso aberto existentes nos repositórios portugueses, e de um serviço de alojamento de repositórios, permitindo a criação de vários novos repositórios em Portugal.

Neste contexto, com o objetivo de promover a qualidade e harmonização dos metadados dos repositórios no Portal RCAAP, foi criado o Validador RCAAP^{iv}. Uma vez que o projeto definiu como obrigatória a implementação das diretrizes DRIVER^v nos repositórios integrados no RCAAP, esta ferramenta, disponível de forma livre para a comunidade, veio permitir uma forma simples e rápida de validação e análise do conteúdo assim como na identificação de erros nos metadados com base nos critérios pré-definidos. Posteriormente, este processo de validação foi integrado com o procedimento de agregação no portal, permitindo que no decurso da atualização diária dos recursos no Portal RCAAP, os erros fossem identificados e os gestores de repositórios recebessem avisos sempre que algum registo não estivesse conforme os critérios definidos.

Tendo por base o trabalho desenvolvido ao longo dos últimos cinco anos e a experiência de utilização desta primeira versão do Validador, e com o objetivo de permitir maior flexibilidade e ampliar o âmbito das validações, o projeto RCAAP apostou em reformular a sua ferramenta de validação.

Esta apresentação pretende destacar os desenvolvimentos desta segunda versão do validador RCAAP, enfatizando a importância dos padrões e normas de interoperabilidade, dos serviços de apoio à melhoria da qualidade dos metadados e destacando os atributos técnicos desta ferramenta de validação de repositórios.

Esta nova versão do serviço vem introduzir novos contextos de validação, adicionar a capacidade de validação de uma comunidade ou coleção e permitir criar novos tipos de validação personalizados. Além destas funcionalidades adicionais, o novo validador incorpora uma ferramenta para validação dos formatos dos ficheiros nos repositórios, orientada à caracterização desses ficheiros e à definição de ações de preservação digital (ferramenta *FITS*).

Uma das principais novidades desta segunda versão do validador é a possibilidade de seleção de qualquer conjunto (*set*) do repositório ou revista (e não apenas o set DRIVER como na versão anterior). Desta forma, será possível por exemplo validar no repositório apenas os registos de uma faculdade, escola, departamento, uma coleção de teses e dissertações, ou no caso das revistas, uma determinada secção. Uma outra significativa alteração corresponde aos vários perfis de validação que podem ser configurados para as futuras necessidades do projeto. Estão disponíveis perfis de validação para as teses e dissertações, para a identificação de campos vazios, para a listagem de registos embargados assim como para os trabalhos financiados pela FCT ou Comissão Europeia.

O processo de validação aplica três análises distintas para cada um dos elementos de metadados:

- Verifica a existência de conteúdo (por exemplo, se tem um título);
- Verifica se o conteúdo está bem formatado (por exemplo, a data);
- Verifica se o conteúdo está de acordo com taxonomias conhecidas (por exemplo, os tipos de documento).

Este novo validador teve substanciais melhorias ao nível do interface, baseado na *framework Bootstrap*, que permite, por exemplo, a compatibilidade com dispositivos móveis.

Na relação com o utilizador, foi acrescentada uma página de ajuda que descreve e apoia a resolução dos problemas identificados. Além disso, e definindo-se como uma ferramenta de uso global, tem uma versão com idioma em inglês para permitir que outros gestores de repositórios e revistas possam usar o validador.

Os conceitos básicos do serviço, tais como a sua disponibilização pública e a integração com o Portal RCAAP mantêm-se, assim como o envio dos relatórios de validação, devidamente reformulados e agora mais completos e pormenorizados.

Conclusão

A criação de um repositório institucional deve adequar-se às normas e diretrizes vigentes na

organização dos recursos para ir ao encontro da interoperabilidade do sistema e garantir a qualidade dos registos depositados.

A disponibilização de recursos em acesso aberto em repositórios institucionais deve garantir um mínimo de coerência entre o documento depositado e a sua descrição, descrever corretamente o tipo de documento, o idioma usado, etc. Nesse sentido, a verificação automática dos registos é uma forma de facilitar a monitorização da qualidade dos metadados e da congruência da informação depositada.

Esta atualização do validador RCAAP oferece aos repositórios e revistas novos serviços e ferramentas para promover a qualidade, a harmonização e a interoperabilidade, a nível nacional e internacional.

Só com uma validação sistemática se consegue garantir a qualidade dos metadados dos repositórios institucionais e desse modo integrar iniciativas conjuntas que adotem as mesmas regras de interoperabilidade. Estas ferramentas permitem ainda obter informação estatística sobre o conteúdo do repositório e identificar erros comuns. Com esta análise é possível atuar nos processos para que as regras estejam embebidas no processo de depósito inicial e deixem de ser sempre corrigidas posteriormente.

ⁱ DCMI: <http://dublincore.org/documents/dcmi-terms>

ⁱⁱ Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH): <http://www.openarchives.org/pmh>

ⁱⁱⁱ Projeto RCAAP: <http://projecto.rcaap.pt>

^{iv} Validador: <http://validador.rcaap.pt/validador>

^v Diretrizes DRIVER (PT) http://www.driver-support.eu/documents/DRIVER_Guidelines_v2_Final__PT.pdf