

# **Integração de informação e interoperabilidade: o Repositório Institucional como suporte à gestão da investigação na Universidade de Aveiro**

*Information integration and interoperability: the role of the institutional repository at University of Aveiro research management support*

**Ana Bela Martins**

Serviços de Biblioteca, Informação Documental e Museologia da Universidade de Aveiro, Portugal.

E-mail: [abela@ua.pt](mailto:abela@ua.pt)

**Bella Nolasco**

Serviços de Biblioteca, Informação Documental e Museologia da Universidade de Aveiro, Portugal.

E-mail: [bellanolasco@ua.pt](mailto:bellanolasco@ua.pt)

**Diana Silva**

Serviços de Biblioteca, Informação Documental e Museologia da Universidade de Aveiro, Portugal.

E-mail: [dianasilva@ua.pt](mailto:dianasilva@ua.pt)

## **Resumo**

A estratégia global de gestão de informação da Universidade de Aveiro prevê um conjunto de sistemas interoperáveis entre si e com outros sistemas externos, de forma a permitir uma racionalização de processos, a assegurar uma adequada gestão de informação e a promover a visibilidade da investigação produzida na instituição. O Repositório Institucional da Universidade de Aveiro (RIA) é uma peça central neste âmbito, tendo sido identificados, no decorrer do desenvolvimento do sistema, um conjunto de funcionalidades fundamentais para tornar ágil o processo de auto depósito de publicações por parte dos autores e para permitir a integração de informação com outros sistemas. São descritas algumas destas funcionalidades e destacados os desafios.

**Palavras-chave:** Repositórios. Auto arquivo. Interoperabilidade. Informação científica.

## **Abstract**

The information management strategy of University of Aveiro provides a set of systems interoperable with each other and with other external systems, to allow process simplification, the promotion of appropriate information management and the research visibility. The Institutional Repository of the University of Aveiro (RIA) is a centerpiece in this area and, during the development of the system, the team had identified a set of key features to make agile the process of deposit of publications and to allow integration with other systems. Some of these features are described and highlighted the challenges.

**Keywords:** Repositories. Interoperability. Scientific information

## **Introdução**

Nos dias de hoje grande parte do conhecimento científico encontra-se disponível em diferentes plataformas sendo que a implementação de repositórios de acesso aberto nas instituições de ensino superior tem sido um dos meios mais usados, nos últimos anos, para a gestão e promoção da visibilidade da produção científica desenvolvida nas instituições. Os repositórios institucionais vieram facilitar mecanismos de auto depósito, de gestão da informação científica, bem como a disponibilização dos conteúdos na Web e a integração e interoperabilidade com outros sistemas. Esta última vertente tem sido apontada recentemente (LEWIS et al., 2009; COAR, 2011) como um dos aspetos fundamentais para a criação de valor no seio das instituições ao nível da gestão da produção científica e disseminação da mesma. De facto, podemos considerar que a interoperabilidade dos repositórios com outros sistemas de informação no âmbito da instituição e fora dela pode ser identificada como um dos fatores críticos para o sucesso e para a adesão dos investigadores ao depósito em repositórios científicos de acesso aberto.

## **O Repositório Institucional da Universidade de Aveiro: desenvolvimento e características**

O Repositório Institucional da Universidade de Aveiro (RIA) é um sistema de informação em regime de auto arquivo que armazena, preserva e dá acesso via Web, em regime de acesso aberto, à produção científica da Universidade de Aveiro (UA). Tem como principais objetivos: a) Promover o conhecimento e o aumento do impacto da investigação produzida na UA, reforçando a sua visibilidade; b) Partilhar o conhecimento científico e técnico alcançado na investigação produzida na UA, contribuindo para a geração de novos conhecimentos; c) Facilitar o acesso à produção científica produzida na UA e aos metadados a ela associados, permitindo uma rápida disseminação via Web da mesma; d) Dotar a universidade de ferramentas que permitam analisar e gerir a sua produção científica; e) Ajudar a preservar a memória intelectual desta instituição universitária; f) Concorrer para a promoção da imagem da UA.

O projeto de desenvolvimento deste sistema teve início em Setembro de 2010, tendo ocorrido várias fases ao nível da configuração, desenvolvimento e gestão da informação. Baseado no software *open source* DSpace, o desenvolvimento e configuração do sistema

esteve a cargo de uma equipa com elementos dos Serviços de Biblioteca, Informação Documental e Museologia da UA e dos Serviços de Tecnologias de Informação e Comunicação da UA. Em todo processo de desenvolvimento e constituição do RIA a equipa beneficiou de apoio por parte do projeto Repositório Científico de Acesso Aberto de Portugal (RCAAP) e da equipa responsável pelo RepositoriUM (Universidade do Minho) ao nível do acesso a tutoriais previamente realizados, exemplos de formulários de metadados DSpace, entre outros. Entre outubro e dezembro de 2011 decorreu a fase de testes para o processo de auto depósito que incidiu nos departamentos de Matemática, Química e Comunicação e Arte da universidade. Em janeiro de 2012 o RIA foi disponibilizado e divulgado a toda a comunidade de autores da UA para auto depósito de publicações.

Ao longo das várias fases de trabalho de desenvolvimento foram identificadas funcionalidades a aplicar ao sistema, que vieram em grande parte a ser implementadas, resultando em *add-ons* ao *software* de origem DSpace 1.6.2. Algumas destas novas configurações implementadas no sistema são resultantes da instalação de *add-ons* partilhados pela equipa de configuração do RCAAP, sistema gerido pela Fundação para a Computação Científica Nacional (FCCN) e pela Universidade do Minho. Parte destas configurações resultaram de desenvolvimento levado a cabo pela equipa da UA. Na figura 1 pode ver-se o diagrama representativo dos principais *add-ons* instalados no RIA.

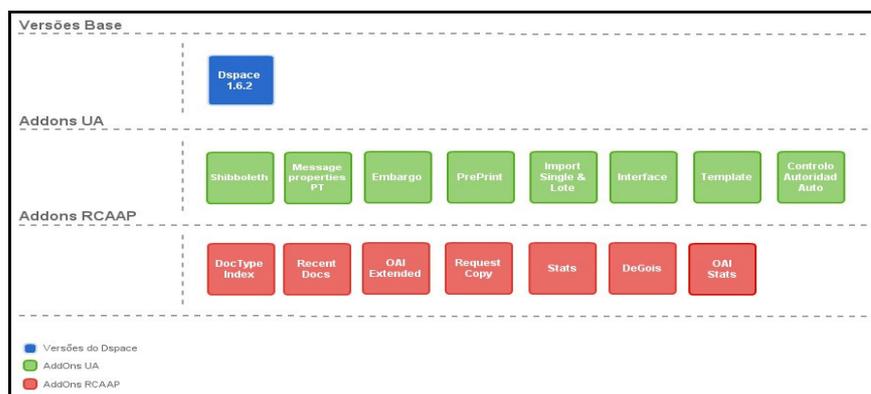


Figura 1: Diagrama de *add-ons* instalados no Repositório Institucional da Universidade de Aveiro.  
 Fonte: Relatório do Repositório Institucional da Universidade de Aveiro - <http://ria.ua.pt> (não publicado)

A gestão de informação no RIA está a cargo da equipa dos Serviços de Biblioteca, Informação Documental e Museologia da UA (SBIDM). As atividades desenvolvidas abrangem todas as vertentes de gestão de informação e organização de dados no sistema, tais como a definição de formulários de dados bibliográficos para cada uma das tipologias

possíveis de documentos a incluir no repositório, a definição da estrutura e hierarquia de comunidades, a gestão de utilizadores e outras. No que diz respeito ao auto depósito de publicações foram concebidas duas opções, sendo que para ambas, o primeiro passo é sempre a autenticação no sistema através das credenciais de Utilizador Universal da UA. As opções são: a) preenchimento do formulário de recolha de dados descritivos do documento e depósito do ficheiro digital. Só após a validação dos dados bibliográficos pela equipa de validação, o registo do documento fica disponível; b) carregamento automático de registos em lote através da importação de ficheiros com dados bibliográficos oriundos de bases de dados na Web (*Web of Science* e SCOPUS) ou de ficheiros em formato RIS.

### **Integração de informação com outros sistemas e controlo de qualidade**

O trabalho de desenvolvimento do RIA está subjacente a uma estratégia global de gestão da informação na UA. A configuração do sistema e a adequação de funcionalidades tem sido promovida em função de uma gestão integrada da informação relativa à produção científica da instituição em termos de avaliação, acesso, visibilidade e preservação, atendendo também à integração e interoperabilidade com plataformas de gestão e acesso à informação científica ao nível nacional e internacional. De facto, a equipe de gestão e desenvolvimento do repositório considerou desde o início como desafios fundamentais a comunicação de informação com sistemas internos e externos de forma a partilhar e integrar informação relevante e a racionalizar recursos e serviços. Alguns destes aspetos são abordados no relatório “The Case for Interoperability for Open Access Repositories version 1.0” (COAR, 2011) colocando em foco questões como os desafios atuais ligados à consistência na identificação de autoria e terminologias, a sustentabilidade dos *standards* e protocolos para normalização bibliográfica e troca de dados e a compatibilidade entre sistemas de informação. Estes e outros desafios têm sido identificados no decorrer do processo de desenvolvimento do RIA e são apresentados de seguida.

### a) Integração RCAAP – protocolo OAI-PMH

A instalação do protocolo OAI-PMH<sup>1</sup> permite disponibilizar a informação descritiva do repositório de forma a que os metadados possam ser agregados por serviços externos. A agregação RCAAP dos registos disponíveis no RIA em acesso aberto é disso exemplo e tem sido fundamental para a visibilidade da informação científica produzida na Instituição. De destacar neste âmbito a aplicação das directrizes DRIVER<sup>2</sup> como forma de normalização de metadados, condição essencial para a integração com este serviço agregador.

### b) Digital Object Identifier

Um dos aspetos críticos essenciais na interoperabilidade entre sistemas passa por se pensar um sistema com um alto nível de granularidade de informação descritiva de forma a “alimentar” outros sistemas com essa informação. A inclusão de mecanismos de identificação inequívoca de registos e documentos digitais através de identificadores persistentes como o *handle* e DOI (Digital Object Identifier) em repositórios afigura-se essencial e tem sido no âmbito do RIA um aspeto tido em conta no processo de validação de registos depositados. Tal como refere Sayão (2012, p. 80)

O sucesso atual dos sistemas distribuídos de informação, tal como a World Wide Web, bem como o potencial de desenvolvimento futuro desses sistemas, depende fortemente da consistência e estabilidade dos links entre recursos digitais que eles disponibilizam online, via rede de computadores. A identificação persistente é uma característica chave para os repositórios digitais - sejam eles bibliotecas ou arquivos - que precisam funcionar segundo os paradigmas de interoperabilidade.

Neste sentido destacamos uma das orientações internas para o processo de validação de registos a disponibilizar online no repositório: a inclusão do DOI relativo ao documento digital nos metadados do registo, como demonstrado na figura 2.

---

<sup>1</sup> Open Archives Initiative Protocol for Metadata Harvesting <http://www.openarchives.org/pmh/>

<sup>2</sup> Diretrizes DRIVER - Digital Repositories Infrastructure Vision for European Research [http://www.driver-support.eu/documents/DRIVER\\_Guidelines\\_v2\\_Final\\_PT.pdf](http://www.driver-support.eu/documents/DRIVER_Guidelines_v2_Final_PT.pdf)

<b>título:</b>	Deeper insight into the monoterpenic composition of <i>Ferula gummosa</i> oleo-gum-resin from Iran
<b>autor:</b>	Jalali, Hossein T. Petronilho, Sílvia Villaverde, Juan José Coimbra, Manuel A. Domingues, M. Rosário M. Ebrahimi, Zahra J. Silvestre, Armando J. D. Rocha, Sílvia M.
<b>palavras-chave:</b>	<i>Ferula gummosa</i> Boiss Oleo-gum-resin Essential oil Monoterpenic composition Comprehensive two-dimensional gas chromatography (GC × GC-ToFMS)
<b>data de publicação:</b>	2012
<b>editor:</b>	Elsevier
<b>resumo:</b>	The monoterpenic composition of the essential oil of oleo-gum-resin <i>Ferula gummosa</i> Boiss from Iran was studied by GC × GC-ToFMS. This study allowed to tentatively identify 130 monoterpenic compounds including 24 hydrocarbons, 52 alcohols, 4 ethers, 20 aldehydes, 8 oxides, 18 ketones, 1 furan and 3 epoxides, from which only 25 were reported before in this species. Furthermore, while previous studies identified terpenic hydrocarbons such as beta- and alpha-pinene as the most abundant monoterpenes, in the present study monoterpene alcohols as myrtenol, mentha-1,8-dien-7-ol, two pinocarvedol isomers and mentha-1,5-dien-8-ol were the predominant components. GC × GC-ToFMS analysis allowed achieving a deeper characterization of the monoterpenic composition of <i>F. gummosa</i> , which may contribute to the increase of the value of this biomass plant material as a source of monoterpenic compounds. Moreover, this approach allows a more efficient quality control for existent applications and to prevent adulterations, promoting market confidence. (C) 2011 Elsevier B.V. All rights reserved.
<b>URI:</b>	<a href="http://hdl.handle.net/10773/9019">http://hdl.handle.net/10773/9019</a>
<b>ISSN:</b>	0926-6690
<b>versão do editor/DOI:</b>	<a href="http://dx.doi.org/10.1016/j.indcrop.2011.11.001">http://dx.doi.org/10.1016/j.indcrop.2011.11.001</a>
<b>fonte:</b>	Industrial Crops and Products
<b>aparece nas coleções:</b>	CICECO - Artigos

Figura 2: Registo bibliográfico RIA.

Fonte: Repositório Institucional da Universidade de Aveiro (<http://ria.ua.pt>)

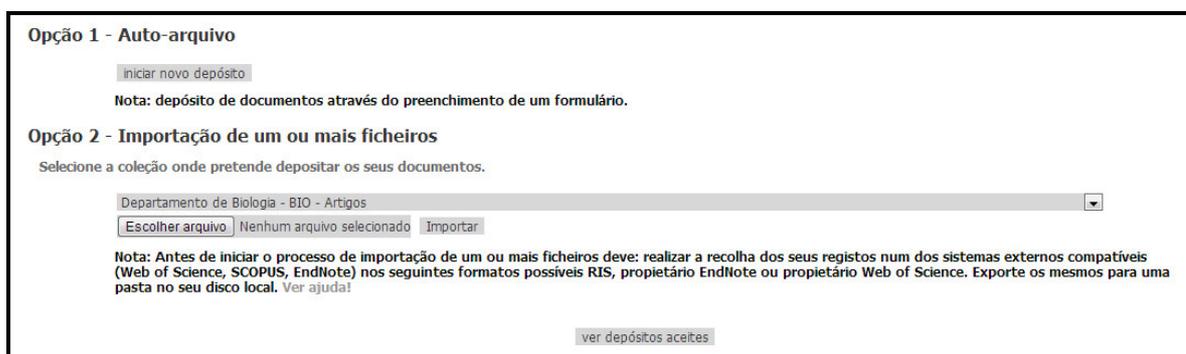
### c) Identificação inequívoca de autoria

A identificação de autoria tem sido amplamente referida como um desafio fundamental para a consistência dos sistemas de informação, tanto ao nível da recuperação de informação pelos utilizadores como no que se refere à integração de informação com outras plataformas. Desta forma, com base numa reflexão apurada da equipa de desenvolvimento do sistema e tendo em conta a análise realizada a outros sistemas de informação da universidade, foi desenvolvido um *add-on* para a integração do RIA com um sistema de informação que realiza a gestão de utilizadores em toda a universidade denominado Registo Central de Utilizadores da UA que inclui docentes, alunos, investigadores e funcionários da instituição. Esta funcionalidade permite a identificação inequívoca dos autores dos documentos no repositório, apresentando inúmeras vantagens, entre as quais a disponibilização de dados de forma controlada para outros sistemas de informação internos, o que resulta numa gestão de dados de investigação e docência de forma segura e controlada.

### d) Carregamento automático de registos em lote através da importação de ficheiros

A possibilidade que os autores dispõem de realizar uma recolha automática de dados bibliográficos de bases de dados na Web ou de sistemas de gestão bibliográfica no momento de depósito no RIA tem concorrido para a simplificação do processo de auto arquivo contribuindo para o seu incentivo. O *add-on* “import single & lote” permite que o autor, após

autenticação no sistema, realize a importação de um ficheiro previamente exportado em formato RIS de uma base de dados como a Web of Science ou a SCOPUS ou de um sistema de gestão de referências bibliográficas. Os dados bibliográficos destes registos importados são transformados em registos no repositório, nesse momento o autor deverá terminar o *workflow* de auto depósito para cada um dos registos, anexando o documento digital correspondente. Esta possibilidade resulta, de uma forma muito efetiva, num enriquecimento do registo bibliográfico com o preenchimento automático de campos descritivos, o que claramente constitui valor acrescentado para o sistema de informação.



The screenshot displays a web interface for document deposition. It features two main sections: 'Opção 1 - Auto-arquivo' with a button 'iniciar novo depósito' and a note about depositing documents via a form; and 'Opção 2 - Importação de um ou mais ficheiros' which includes a dropdown menu for selecting a collection (currently set to 'Departamento de Biologia - BIO - Artigos'), buttons for 'Escolher arquivo', 'Nenhum arquivo selecionado', and 'Importar', and a detailed note about compatible external systems (Web of Science, SCOPUS, EndNote) and file formats (RIS, etc.). A 'ver depósitos aceites' button is located at the bottom right of the interface.

Figura 3: Opções de auto arquivo de publicações no RIA.

Fonte: Repositório Institucional da Universidade de Aveiro (<http://ria.ua.pt>)

### e) Integração de informação com o sistema Gestão de Informação para Avaliação de Docentes da UA

O sistema PADUA – Gestão de Informação para Avaliação de Docentes trata-se de uma plataforma desenvolvida na UA que permite gerir de uma forma integrada o processo de avaliação docente desta instituição. Trata-se de um sistema desenvolvido internamente com ligação à base de dados de utilizadores da Instituição e funciona com auto submissão de informação por parte dos docentes. Um dos itens considerados neste sistema para a avaliação de docentes refere-se à produção científica, considerando uma lista de tipologias de publicações aceites para inclusão na plataforma, assumindo uma cotação diferente para cada uma das tipologias aceites. No sentido de facilitar o processo de submissão de informação relativa às publicações científicas, permitindo aos autores a realização desta tarefa apenas uma vez, foi desenvolvido um *Web service* que permite a disponibilização de informação do RIA para o PADUA. A comunicação entre os dois sistemas realiza-se por intermédio de um *Web service* que permite ao docente autenticado no PADUA efetuar a recolha de informação disponibilizada no RIA, selecionando os registos no repositório a incluir no sistema de

avaliação. Por outro lado, de referir que a troca de dados entre os dois sistemas só é permitida por meio da identificação comum da autoria, possível mediante a ligação de ambos os sistemas ao Registo Central de Utilizadores da UA, que atribui a cada utilizador um código único, o que permite a sustentabilidade desta função. Do ponto de vista institucional observa-se como muito vantajosa a possibilidade de os autores poderem executar a função de depósito de informação numa única plataforma, o que vem poupar tempo, criar uma homogeneidade nos processos, o que obviamente muito concorre para a motivação dos autores para estas atividades. Por outro lado, os registos, após o processo de validação no repositório, poderão ser “distribuídos” por outros sistemas internos. A este propósito, Joint (2008, p. 573) aponta a importância de um único ponto de acesso para a inserção de registos:

Either research is self-archived onto a research publications system and then pushed to the open access repository, or open access deposits are pushed to the research publications database. As long as duplication of effort is avoided, the direction of flow is irrelevant. But it is essential to have a single point of upload into a master repository where editing and quality control takes place, before reliable records are fed out to other parts of the campus research information system.

A integração de informação entre o RIA e o PADUA revelou-se profícua e de grande importância no contexto da estratégia de gestão de informação na universidade. Perspetiva-se no âmbito desta estratégia outros projetos que permitam uma gestão global e adequada da informação relativa à investigação produzida na UA.

### **Desafios futuros**

O projeto de criação de um sistema único para a gestão da investigação da universidade e preparado para efetuar automaticamente a recolha de dados bibliográficos no repositório institucional constitui-se como uma das peças fundamentais da estratégia de gestão da informação ao nível da produção científica na UA. Tal como apontam Jeffery e Asserson (2009, p. 74) “For the end-user access to heterogeneous distributed CRIS and repositories should appear homogeneous and local to the end-user”. Em fase de diagnóstico perspectiva-se a necessidade de uma recolha exaustiva dos vários sistemas em uso pelas unidades de investigação para a gestão da informação relativa a projetos de investigação, produtividade científica e outros. Nesta fase inicial é essencial uma comunicação ágil com os responsáveis pelas unidades de investigação, de forma a serem definidos os vários critérios e consideradas as especificidades de cada uma. A racionalização de processos, a promoção de uma adequada

gestão de informação e a promoção da visibilidade da investigação produzida na universidade são os objetivos essenciais deste projeto.

## Conclusão

No contexto universitário demonstra-se essencial uma gestão integrada da informação relativa à produção científica da instituição em termos de avaliação, acesso, visibilidade, preservação, atendendo também à integração e interoperabilidade com plataformas de gestão e acesso à informação científica ao nível nacional e internacional. São vários os desafios que se colocam neste âmbito considerando que a interoperabilidade entre sistemas e a integração de informação bibliográfica de uma forma válida e controlada depende da aplicação de protocolos, como o OAI-PMH, do uso de *standards* e *guidelines* relativos aos esquemas de metadados e à descrição bibliográfica, da consistência na identificação de autoria e terminologias. Estes e outros desafios têm sido identificados no decorrer do processo de desenvolvimento do RIA, tendo sido desenvolvidos um conjunto de *add-ons* que facilitam os processos. De referir ainda que a reflexão comum em torno destes aspetos será um meio de potenciar o desenvolvimento de soluções adequadas a nível global.

## Referências

COAR. **The case for interoperability for open access repositories:** version 1.0. July 2011. 10 p.

CORNELL UNIVERSITY LIBRARY (org.). **Open Archives Initiative Protocol for Metadata Harvesting.** Disponível em: < <http://www.openarchives.org/pmh/>>. Acesso em: 2 de maio de 2013.

DRIVER; SDUM (trad.). **Directrizes DRIVER 2.0: directrizes para fornecedores de conteúdos, exposição de recursos textuais com o protocolo OAI-PMH:** DRIVER - Digital Repositories Infrastructure Vision for European Research 144 p. 2009. Disponível em: < [http://www.driver-support.eu/documents/DRIVER\\_Guidelines\\_v2\\_Final\\_PT.pdf](http://www.driver-support.eu/documents/DRIVER_Guidelines_v2_Final_PT.pdf)>. Acesso em: 11 de março de 2013.

JEFFERY, K.; ASSERSON, A. Institutional repositories and current research information systems. **New Review of Information Networking**, v. 14, n. 2, p. 71-83, 2009. Disponível em: < <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=45537723&lang=pt-br&site=eds-live>>. Acesso em: 26 de abril de 2013.

JOINT, N. Current research information systems, open access repositories and libraries: ANTAEUS. **Library Review**, v. 57, n. 8, p. 570-575, 2008. Disponível em: <<http://dx.doi.org/10.1108/00242530810899559>>. Acesso em: 8 de abril de 2013.

LEWIS, S. et al. If SWORD is the answer, what is the question?: use of the simple web-service offering repository deposit protocol. **Program: electronic library and information systems**, v. 43, n. 4, p. 407-418, 2009. Disponível em: <<http://dx.doi.org/10.1108/00330330910998057>>. Acesso em: 22 de março de 2013.

SAYÃO, L. F. Interoperabilidade das bibliotecas digitais: o papel dos sistemas de identificadores persistentes-URN, PURL, DOI, Handle System, CrossRef e OpenURL. **Transinformação**, v. 19, n. 1, 2012.

**DOI:** [10.11606/issn.2178-2075.v4i2p202-211](http://dx.doi.org/10.11606/issn.2178-2075.v4i2p202-211)