

Integração do mecanismo de sincronização PTCRISync com o CRIS do ISCTE-IUL

Integration of the PTCRISync synchronization mechanism with ISCTE-IUL's CRIS

Integración del mecanismo de sincronización PTCRISync con el CRIS del ISCTE-IUL

António Luís Lopes | alsl@iscte-iul.pt

Instituto Universitário de Lisboa. Lisboa, Portugal

Resumo

A inovação e conhecimento gerados pela investigação científica só cumprem o seu objetivo se forem largamente difundidos e estiverem acessíveis à comunidade científica em particular e ao público em geral. Contudo, a multiplicidade de sistemas de gestão de ciência e repositórios institucionais existentes não permite a simples e rápida divulgação dos resultados das atividades científicas. Sob a máxima “Introduza uma vez, reutilize múltiplas”, o programa PTCRIS foi criado em 2014 com o objetivo de criar um ecossistema integrado de gestão de informação científica. Este ecossistema, que inclui o desenvolvimento da framework PTCRISync, permite que diferentes sistemas de informação possam partilhar e sincronizar informação entre eles evitando assim a duplicação de esforços dos investigadores ao não terem de depositar a mesma informação em diversos sistemas diferentes. Este artigo descreve a integração do mecanismo de sincronização PTCRISync no sistema de gestão de informação científica do ISCTE-Instituto Universitário de Lisboa, o Ciência-IUL. São também apresentados os resultados e o impacto da implementação referida em termos de benefícios para os investigadores e instituições.

Palavras-chave: Sistemas de Informação de Investigação Atual; Interoperabilidade entre CRIS; Sincronização; PTCRISync; Ciência-IUL.

Abstract

The innovation and knowledge generated by scientific research will only fulfil its goal if they are largely propagated and accessible to the scientific community in particular and society in general. However, the multiplicity of existing science information systems and institutional repositories does not allow for a simple and fast propagation of the results of scientific activities. Under the dictum “Add once, reuse multiple times”, the PTCRIS initiative was started in 2014 with the goal of creating a scientific information integrated ecosystem. This ecosystem, which includes the development of the PTCRSync *framework*, allows for different information systems to share and synchronize information amongst them thus avoiding the duplication of efforts of the researchers by not having to deposit the same information in several different systems. This paper describes the integration of the PTCRSync synchronization mechanism with the current research information system of ISCTE-Instituto Universitário de Lisboa, Ciência-IUL. The results and the impact of the referred implementation in terms of the benefits for researchers and institutions are also presented.

Keywords: Current research information systems; Interoperability between CRIS; Synchronization; PTCRISyn; Ciência-IUL.

Resumen

La innovación y el conocimiento generados por la investigación científica sólo cumplen su objetivo si son ampliamente difundidos y están accesibles a la comunidad científica en particular y al público en general. Sin embargo, la multiplicidad de sistemas de gestión de ciencia y repositorios institucionales existentes no permite la simple y rápida divulgación de los resultados de las actividades científicas. Bajo la máxima “Introduzca una vez, reutilice múltiples”, el programa PTCRIS se creó en 2014 con el objetivo de crear un ecosistema integrado de gestión de información científica. Este ecosistema, que incluye el desarrollo del marco PTCRSync, permite que diferentes sistemas de información puedan compartir y sincronizar información entre ellos evitando así la duplicación de esfuerzos de los investigadores al no tener que depositar la misma información en diversos sistemas diferentes. Este artículo describe la integración del mecanismo de sincronización PTCRSync en el sistema de gestión de información científica del ISCTE-Instituto Universitário de Lisboa, Ciência-IUL. También se presentan los resultados y el impacto de la aplicación referida en términos de beneficios para los investigadores e instituciones.

Palabras-clave: Sistemas de información de investigación actual; Interoperabilidad entre CRIS; Sincronización; PTCRISync; Ciência-IUL.

Introdução

A multiplicidade de sistemas de informação relacionados com a gestão dos resultados da investigação científica dificulta a tarefa de divulgação e partilha da inovação e conhecimento gerado nas atividades científicas. Vejamos o caso de um investigador que pertença a um dos centros de investigação do ISCTE-IUL. Neste caso, o investigador é responsável por colocar as suas produções científicas no sistema de gestão científica, o Ciência-IUL (<https://ciencia.iscte-iul.pt/>), por este estar ligado ao sistema de avaliação de desempenho. Por outro lado, também deverá colocar essas mesmas produções científicas no repositório institucional (para divulgação externa). Dependendo do centro de investigação a que pertence, poderá ainda também ter de inserir a mesma informação no *backoffice* do portal do respetivo centro de investigação. Para além destas responsabilidades no contexto da sua própria instituição, o investigador tem também de manter currículos atualizados em sistemas externos, como forma de divulgação científica ou de obtenção de financiamento para projetos de investigação. A consequência evidente deste tipo de abordagem na divulgação científica é que uma parte considerável do tempo de um investigador seja passado a inserir informação em sistemas e formatos diferentes (com variável dificuldade ou usabilidade).

Integração entre sistemas locais no ISCTE-IUL

Esta preocupação levou a que nos últimos anos se desenvolvesse, no ISCTE-IUL, um esforço de integração entre os vários sistemas de informação de forma a que os investigadores só tenham de inserir a informação uma única vez e que esta seja depois propagada pelos restantes sistemas sem haver necessidade de intervenção extra por parte dos investigadores. Atualmente, os investigadores só inserem a informação no Ciência-IUL e este é depois responsável por propagar a informação para o sistema de avaliação de desempenho, para o repositório institucional e para os portais dos diversos centros de investigação da instituição. Esta abordagem apresentou ganhos evidentes e significativos na vida dos investigadores, por lhes permitir otimizar o seu tempo dedicado à investigação. Também para o repositório institucional, a integração com o Ciência-IUL contribuiu de forma muito significativa para aumentar o depósito de artigos científicos publicados em revistas nacionais/internacionais com revisão por pares, disponibilizando assim aos utilizadores conteúdo científico relevante nas áreas de conhecimento da instituição¹⁻².

O panorama global

Apesar deste esforço de integração entre sistemas do ISCTE-IUL ter permitido atingir um novo patamar de desenvolvimento e crescimento na preservação e disseminação da produção científica da instituição, os investigadores têm também responsabilidades de manter informação de índole científica atualizada noutros sistemas externos à instituição, nomeadamente outros sistemas de gestão científica nacionais e internacionais, sistemas de curriculum científico e plataformas de submissão de propostas para obtenção de financiamento científico. Da mesma maneira que anteriormente no ISCTE-IUL, os investigadores eram afetados pela multiplicidade de sistemas em que tinham de colocar informação repetida, também no panorama nacional e internacional, os investigadores são confrontados com uma grande heterogeneidade de sistemas e formatos em que a informação científica é preenchida e que leva a que os dados estejam normalmente desatualizados, inconsistentes ou mesmo ausentes.

A iniciativa PTCRIS e a *framework* PTCRISync

Neste sentido, e obedecendo à máxima “Introduza uma vez, reutilize múltiplas”, a unidade FCCN (Fundação para a Computação Científica Nacional) da Fundação para a Ciência e Tecnologia (FCT) desenvolveu a iniciativa PTCRIS³ com o intuito de criar um ecossistema que liga os diversos sistemas de gestão científica nacionais (e de certa maneira internacionais) de forma a que o investigador só tenha que inserir a informação num deles e que depois seja possível sincronizar a informação com todos os restantes sistemas. Para cumprir esta visão, o PTCRIS tem dois objetivos principais. O primeiro consiste em definir uma *framework* regulatória baseada nas melhores práticas e *standards* internacionais no que toca à representação e partilha de informação de índole científica. O segundo objetivo consiste em fomentar a adoção desta *framework* nos diversos sistemas de informação, tanto de âmbito nacional como de âmbito local (ao nível de instituições individuais).

Um dos resultados principais desta iniciativa é a *framework* de sincronização PTCRISync⁴, um mecanismo uniforme e modular de sincronizar produções científicas entre cada sistema local e um *hub* central do ecossistema, neste caso o ORCID (<https://orcid.org/>). Num cenário típico da utilização deste mecanismo de sincronização, um investigador poderá exportar as publicações que tem num dos sistemas locais para o ORCID (deixando-as assim disponíveis para serem recolhidas por outros sistemas membros do ecossistema) assim como importar publicações que tenham sido adicionadas ao ORCID por outro sistema. Nesse sentido, a *framework* PTCRISync especifica não só os métodos que todos os membros do ecossistema devem utilizar para garantir a sincronização, como também especifica o algoritmo que define

a ordem e a forma como esses métodos devem ser invocados. Para facilitar a integração dos sistemas locais com a *framework*, foi desenvolvida uma biblioteca de software Java (o código fonte da biblioteca encontra-se disponível no GitHub: <https://github.com/fccn/PTCRISync>) de acordo com a especificação do algoritmo de sincronização. Esta biblioteca de software foi usada diretamente no Ciência-IUL (uma vez que este também foi desenvolvido na linguagem de programação Java) para permitir desencadear o mecanismo de sincronização com o ORCID. Está em desenvolvimento também uma versão SaaS (*Software-as-a-Service*) do mecanismo de sincronização PTCRISync para que outros sistemas possam integrar-se no ecossistema PTCRIS sem que tenham de usar a biblioteca de software desenvolvida em Java.

Princípios gerais de sincronização

Para que um sistema de informação de uma instituição possa fazer parte do ecossistema PTCRIS tem de obedecer aos princípios estabelecidos na *framework* PTCRISync, para garantir a consistência da sincronização entre os vários membros do ecossistema. Um dos princípios fundamentais passa pela forma como o sistema interage com o utilizador para realizar o processo de sincronização. Cada sistema terá de ser capaz de recolher a informação necessária do lado do utilizador (nomeadamente o seu identificador ORCID e quais as publicações que deseja importar/exportar de/para o ORCID) e posteriormente desencadear o processo de sincronização de forma a que seja garantida a consistência de informação entre os vários sistemas e sejam apresentadas ao investigador quais as decisões que há a tomar face ao resultado da sincronização desencadeada.

Outro dos princípios fundamentais aborda a forma como o sistema deve realizar os dois principais passos no mecanismo de sincronização: a exportação e a importação. A exportação (que consiste no envio da informação marcada como sincronizada no sistema para o ORCID) deverá ser desencadeada antes da importação, para garantir a consistência da informação sincronizada, isto é, para garantir que informação nova que tenha sido adicionada/editada do lado do sistema esteja refletida do lado do ORCID. Assim, quando em seguida se executa a importação, já sabemos que a informação que é recolhida do lado do ORCID já contém os dados mais atualizados, diminuindo assim a possibilidade de ocorrência de inconsistências de informação.

Adoção da *framework* PTCRISync no ISCTE-IUL

No contexto do projeto piloto PTCRISync (projeto que conta com a participação da FCT|FCCN, Universidade do Minho, ISCTE-IUL e Universidade de Évora), o ISCTE-IUL integrou o seu sistema de gestão de produção científica, o Ciência-IUL, com o ecossistema PTCRIS. Esta integração permite a todos os docentes e investigadores da instituição sincronizarem a sua produção científica com o sistema ORCID automaticamente e conseqüentemente com qualquer outro sistema que passe a fazer parte do ecossistema.

De forma a permitir a integração do Ciência-IUL com o ecossistema PTCRIS, houve a necessidade de realizar algumas adaptações no sistema. Em primeiro lugar, foi criada uma secção específica no *backoffice* dos investigadores dedicada exclusivamente à sincronização com o ORCID (através da *framework* PTCRISync). Uma vez que o ecossistema PTCRIS também inclui uma série de diretrizes que estabelecem a forma como a interação com o utilizador é realizada (não só ao nível aplicacional, mas também de interface visual e usabilidade), a integração direta com a atual secção de gestão de publicações não era possível. Assim sendo, a criação da secção específica para lidar com a parte de sincronização permitiu respeitar por completo as diretrizes definidas na *framework*.

A sincronização através da *framework* PTCRISync exige que determinada informação seja mantida do lado de cada sistema local, nomeadamente o registo de autorização de ligação à conta ORCID e o registo de quais as publicações a sincronizar para cada investigador. Para permitir a monitorização e acompanhamento do sistema, é também recomendado manter um registo de estatísticas de utilização do

mecanismo de sincronização, registrando o número de publicações importadas, exportadas e removidas para cada investigador. Foi necessário, portanto, criar no Ciência-IUL os modelos de dados que permitem armazenar a informação referida.

Antes de dar início ao processo de sincronização entre o Ciência-IUL e o ORCID, levado a cabo pela *framework* PTCRISync, é necessário primeiro estabelecer a ligação entre o perfil do investigador no Ciência-IUL e o seu perfil no ORCID. Este passo não é executado na *framework* PTCRISync e é da responsabilidade de cada sistema local realizar este passo para obter o token de autenticação necessário ao funcionamento do processo de sincronização. No caso do Ciência-IUL, este passo é iniciado com a apresentação do botão para efetuar a ligação à conta ORCID do investigador, como evidenciado na Figura 1.

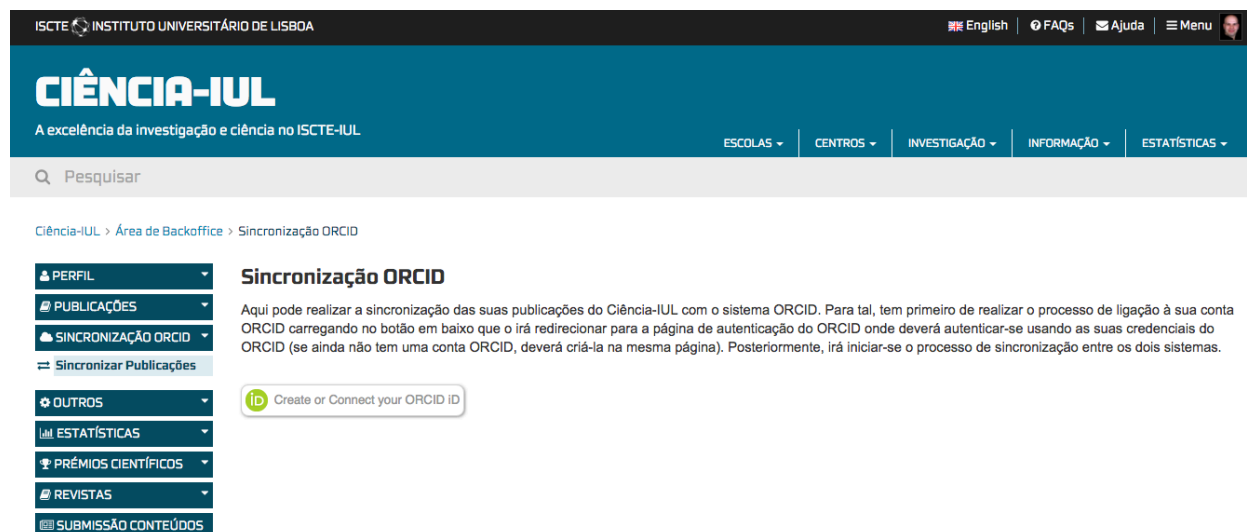


Figura 1 - Botão para ligação à conta ORCID

Ao clicar no botão, o investigador é redirecionado para a página de autenticação do ORCID onde tem a possibilidade de autorizar o acesso à conta (por parte do Ciência-IUL) realizando o respetivo *login*, se já tiver previamente criado uma conta, ou realizando o registo de uma nova conta. Após ser dada a autorização, o utilizador é redirecionado de volta ao Ciência-IUL onde é desencadeado o mecanismo de sincronização entre os dois sistemas e são então apresentadas ao utilizador as diferenças entre os dois perfis (ver Figura 2):

- Publicações que estão no ORCID mas que não existem no Ciência-IUL. Estas aparecem com a sugestão “Adicionar ao Ciência-IUL”
- Publicações que estão no Ciência-IUL mas que não estão no ORCID. Estas aparecem com a sugestão “Sincronizar com o ORCID”

The screenshot shows the 'Sincronização com o ORCID' interface. At the top, there are navigation tabs: 'Por Sincronizar (35)', 'Avisos (0)', and 'Sincronizadas (3)'. Below this, a warning message states: 'Este ecrã apresenta as diferenças entre o seu perfil no Ciência-IUL e o seu perfil no sistema ORCID (sob a forma de notificações). Os registos que existem no Ciência-IUL mas que não existem no ORCID (ou que ainda não estão sincronizados), são apresentados aqui com a opção "Sincronizar com o ORCID". Os registos que existem no ORCID mas que não têm registo correspondente no Ciência-IUL, são apresentados aqui com a opção "Adicionar ao Ciência-IUL". Escolha quais os registos que pretende sincronizar com o sistema ORCID e carregue no botão "Sincronizar". Ao marcar um registo para ser sincronizado, todas as alterações que posteriormente faça a esse registo no Ciência-IUL serão repercutidas no ORCID. Dadas as restrições e limites de consumo do serviço ORCID, a sincronização entre os registos poderá não ocorrer imediatamente. Para mais informações, consulte as FAQs.' Below the warning, there is a table of items to be synchronized:

<input type="checkbox"/> Selecionar Tudo	Registos a sincronizar
<input type="checkbox"/> Adicionar ao Ciência-IUL	<p>Capítulo de livro Efficient algorithms for agent-based semantic resource discovery Agents and Peer-to-Peer Computing, 2012 DOI: 10.10123/12345</p>
<input type="checkbox"/> Sincronizar com o ORCID	<p>Artigo em revista científica Distributed Coordination of Heterogeneous Agents Using a Semantic Overlay Network and a Goal-Directed Graphplan Planner PLoS One, 2013 Lopes, A. L.; Botelho, L.; DOI (fonte: próprio): 10.1371/journal.pone.0062931 ; WoS (fonte: próprio): 000319330200014 ; Handle (fonte: Ciência-IUL): http://hdl.handle.net/10071/5568 ; WoS (fonte: ORCID): WOS:000319330200014 ; WoS (fonte: Ciência-IUL): 000319330200014 ; Scopus (fonte: Ciência-IUL): 2-s2.0-84877934803 ; ID Ciência-IUL: ciencia-iul-10476</p>

Figura 2 - Diferenças entre os dois perfis

Como se pode ver na Figura 2, o utilizador tem a possibilidade de decidir o que fazer em relação às notificações resultantes da análise feita na comparação entre o perfil do Ciência-IUL do investigador e o respetivo perfil ORCID. Após seleção e confirmação por parte do utilizador sobre quais as publicações que pretende sincronizar, o sistema desencadeia novamente o mecanismo de sincronização, desta vez transferindo as respetivas publicações de um sistema para o outro (de acordo com as decisões do utilizador).

Mais sistemas no ecossistema PTCRIS

O Ciência-IUL passou a fazer parte do ecossistema PTCRIS em fevereiro de 2017 e tem, desde essa altura, permitido aos investigadores do ISCTE-IUL poupar imenso tempo na atualização da sua informação de produção científica no ORCID. O próximo sistema que passará a fazer parte do ecossistema é o Ciência Vitae (<https://www.cienciavitae.pt/>). O Ciência Vitae será o novo sistema nacional de gestão curricular de ciência em Portugal e, quando entrar em funcionamento, ainda em 2017, virá já equipado com o mecanismo de sincronização PTCRISync permitindo assim aos investigadores do ISCTE-IUL (e de outras instituições que entretanto adiram ao ecossistema PTCRIS) atualizarem o seu currículo científico neste sistema sem a necessidade de manualmente inserirem as suas publicações.

Resultados

Podemos estimar o impacto temporal e financeiro que esta medida teve no ISCTE-IUL desde fevereiro de 2017 (altura da integração no ecossistema PTCRIS) e que terá no futuro à medida que mais sistemas forem sendo adicionados e mais investigadores comecem a usar a funcionalidade. À data atual (setembro de 2017), mais de 120 investigadores usaram esta funcionalidade para sincronizar mais de 5500 publicações com o ORCID. Se estimarmos que cada publicação demora 3 minutos a ser inserida num sistema de informação, esta funcionalidade já permitiu a esses investigadores pouparem mais de 16500 minutos (275 horas) só na inserção dessa informação no ORCID.

Quanto mais sistemas forem sendo inseridos no ecossistema, maior será o tempo poupado na inserção dessa informação. Se considerarmos que o salário médio anual de um investigador em Portugal é de 29 000 € (o valor de uma hora de trabalho corresponderá, portanto, a cerca de 15€), estes são os ganhos efetivos e potenciais só para os investigadores do ISCTE-IUL (é importante salientar que estes números não tomam em conta a necessidade de o trabalho ter de ser replicado pelos diversos coautores das publicações):

	Ganhos Reais (1 sistema, ORCID)	Ganhos Potenciais (Número de sistemas adicionados ao ecossistema)		
	1	2	3	4
Atualmente (120 investigadores, 5500 publicações)	275 horas 4 125 €	550 horas 8 250 €	825 horas 12 375 €	1100 horas 16 500 €
Futuramente (1200 investigadores, 27000 publicações)	1 350 horas 20 250 €	2 700 horas 40 500 €	4 050 horas 60 750 €	5 400 horas 81 000 €

Os ganhos para a instituição espelhados na tabela (tanto reais como potenciais) são evidentes e podem servir de motivador para que outras instituições se integrem com o ecossistema PTCRIS. Se considerarmos o panorama nacional, o tempo reduzido neste tipo de tarefas poderá representar uma poupança de mais de meio milhão de euros.

Conclusão

O panorama atual de partilha de informação científica destaca-se pela multiplicidade de sistemas de informação e a duplicação de esforços dos investigadores no que toca à inserção de informação sobre publicações e outros outputs científicos.

O projeto PTCRIS, com a sua *framework* de sincronização PTCRISync, pretende mudar esse panorama através da criação de diretrizes e regulamentação baseadas nas melhores práticas e *standards* internacionais no que toca à representação e partilha de informação de índole científica.

A adoção desta *framework* regulatória no sistema de informação científica do ISCTE-IUL provou ser eficaz na poupança de tempo e recursos que poderão agora ser usados para os fins que mais interessam: investigação e produção de ciência.

Com a adoção da *framework* por parte de outros sistemas nacionais e internacionais, estaremos a construir um panorama futuro em que a partilha de informação científica é rápida e eficiente, permitindo assim termos realmente a ciência totalmente aberta ao mundo.

Referências:

1. Amante MJ, Lopes S, Marçal B, Segurado T. A interoperabilidade entre o Repositório e um sistema CRIS: o caso do ISCTE-IUL. *Cadernos BAD*; (2): 83-93. 2014.
2. Amante MJ, Lopes S, Marçal B, Segurado T. Partilhar Ciência: o repositório do ISCTE-IUL e a interoperabilidade com um sistema CRIS. 12º Congresso Nacional BAD. 2015.
3. Santos A, Ramalho JC, Ferreira M, Silva H, Faria L, Moreira JM, et al. PT-CRIS: Um miradouro sobre o universo científico nacional. 12º Congresso Nacional BAD. 2015.
4. Moreira JM, Cunha A, Macedo N. An ORCID based synchronization framework for a national CRIS ecosystem. *F1000Research*; 4:181. 2015.