

## Digital Object Identifier (DOI) viabilizando a via verde

---

Paulo Eduardo Potyguara Coutinho–Marques

Instituto de Comunicação e Informação Científica e Tecnológica em Saúde –  
Fundação Oswaldo Cruz  
pmarques@fiocruz.br

Rachel Magarinos–Torres

Faculdade de Farmácia – Universidade Federal Fluminense  
racheltorres.uff@gmail.com

Maria da Conceição Rodrigues de Carvalho

Instituto de Comunicação e Informação Científica e Tecnológica em Saúde –  
Fundação Oswaldo Cruz  
conceicao.carvalho@icict.fiocruz.br

Marcelo Simão de Vasconcellos

Instituto de Comunicação e Informação Científica e Tecnológica em Saúde –  
Fundação Oswaldo Cruz  
vasconcellos@fiocruz.br

Jesús Mena–Chalco

Centro de Matemática, Computação e Cognição – Universidade Federal do  
ABC  
jmenac@gmail.com

Maria Cristina Soares Guimarães

Instituto de Comunicação e Informação Científica e Tecnológica em Saúde –  
Fundação Oswaldo Cruz  
cristina.guimaraes@icict.fiocruz.br

## Introdução

As Tecnologias de Informação e Comunicação (TICs) contribuem com um amplo leque de estratégias para fortalecimento das iniciativas do Acesso Livre à informação científica. Quando se toma a via verde, processo nomeado para o arquivamento da produção científica pelo próprio autor em ambientes de acesso livre como os repositórios institucionais (RI), muito se discute sobre sua sustentabilidade, especificamente quanto ao povoamento destes RIs.

O autoarquivamento, peça fundamental para a via verde, é a principal meta a ser perseguida, mas como é uma estratégia que traz implicações para além de uma política de gestão de um RI, é fundamental que se busque alternativas para povoamento dos mesmos.

Uma possibilidade de povoamento dos RIs é aberta pelo uso de *Digital Object Identifier (DOI)*, ou, identificadores digitais de objeto, atribuídos a artigos científicos quando publicados em formato eletrônico. O DOI consiste em um *link* persistente que garante sua recuperação e sua integridade ao longo do tempo (Sayão, 2010). Nessa perspectiva, o DOI se apresenta como uma estratégia para promover uma ligação e complementariedade entre a Via Dourada (periódicos eletrônicos de livre acesso) e os RIs, contribuindo assim para seu povoamento (Guédon, 2004). Assim, a disponibilidade de uma fonte de dados de artigos científicos acompanhados dos respectivos DOIs se apresenta como estratégia fundamental a ser explorada para contribuir para o fortalecimento dos RIs.

No Brasil há possibilidade de viabilizar o povoamento dos RIs utilizando o DOI, pois o país conta com um banco de dados secundários, mantido pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Ministério da Ciência e Tecnologia, que reúne o registro da produção científica e técnica dos pesquisadores. O Currículo Lattes é um cadastro nacional de currículos de pesquisadores e, em 2013, já somava um total de 2.750.000 registros. Devido ao volume de dados, do preenchimento manual e auto-declaratório e ao uso de dados semi-estruturados, são inúmeros os desafios computacionais para o acesso e manejo dos dados. Seu uso principal tem sido nas iniciativas de avaliação individual dos pesquisadores, bem como de grupos de pesquisa (Mena-Chalco; Cesar Junior, 2009). O registro do currículo neste banco torna-se obrigatório e crucial para executar atividades inerentes à carreira de pesquisador, pois ele é avaliado por seu currículo nesta plataforma quando submete um projeto para financiamento ou quando concorre a um cargo público. Os programas de pós-graduação de universidades públicas também utilizam o currículo desta base para avaliar seus candidatos. É possível recuperar o DOI presente nas referências da produção científica constantes no Currículo Lattes do pesquisador.

## Objetivo

Este trabalho é parte do financiamento recebido do Programa de Indução à Pesquisa e Desenvolvimento Tecnológico (PIPDT) do Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT) da Fundação Oswaldo Cruz (Fiocruz).

Este estudo teve como objetivo o desenvolvimento de uma metodologia para povoamento do repositório institucional da Fundação Oswaldo Cruz (Fiocruz), uma das principais instituições brasileiras de ensino, pesquisa e desenvolvimento na área das ciências da saúde, por meio da coleta de DOIs dos artigos científicos registrados por seus pesquisadores na Plataforma Currículo Lattes.

## Metodologia

O caminho metodológico eleito envolveu definição da amostra. Recorreu-se à Assessoria de Pesquisa que informou uma lista de pesquisadores que sabidamente possui produção científica em quantidade e qualidade de acordo com a qualificação *qualis* CAPES. A partir desta lista, foi coletado o número de identificação de cada pesquisador no Currículo Lattes. Este número de identificação, conhecido como ID Lattes, é único para cada pesquisador que possui seu CV cadastrado no Lattes e é constituído por dezesseis algarismos. Esta lista de pesquisadores com o ID Lattes foi inserida em um serviço na internet desenvolvido exclusivamente para este trabalho. Esta ferramenta recebeu o nome de Web ScriptLattes.

O Web ScriptLattes (WSL) é uma interface visual web que interage com o ScriptLattes (SL). O SL foi produto de uma pesquisa (MENA-CHALCO; CESAR JUNIOR, 2009). O WSL foi desenvolvido no *framework* WEB2PY utilizando a linguagem de programação Python e, dentre suas funcionalidades, encontram-se: (i) interfacear o usuário com o ScriptLattes, (ii) gerar planilha contendo a produção científica revisada por pares com sua respectiva qualificação *qualis* CAPES e (iii) gerar estrutura de dados para importação no Repositório Institucional.

A partir da coleta dos DOIs existentes na lista da produção científica do pesquisador no CV Lattes, foi verificado se o *link* persistente pertencia à algum dos periódicos suportados pelo ScieloBR. O ScieloBR foi escolhido por, além de ser composto por um conjunto de periódicos científicos de reconhecida qualidade, possuir uma interface visual padrão para todos os periódicos científicos disponibilizados por ele. Também foi realizado a correlação entre os metadados oferecidos pelo ScieloBR e o padrão *Dublin Core* utilizado pelo Repositório Institucional. Estando o artigo em algum dos periódicos pertencentes ao ScieloBR, os metadados do artigo assim como o próprio objeto digital eram recuperados e, observando as características da estrutura de importação/exportação de itens da ferramenta *DSpace* utilizada para o Repositório Institucional, foi criada a referida estrutura para

importação já convertendo os metadados utilizados pelo ScieloBR pelos utilizados pelo RI. Os artigos foram importados em um repositório de certificação para uma única coleção. Após esta importação, profissionais de informação verificaram os itens importados e os moveram para coleções idênticas às existentes no RI. Após sinalização pelos profissionais de informação, os itens eram migrados do repositório de certificação para o RI.

## Resultado e conclusão

Como resultado do trabalho desenvolvido apresenta-se o Web ScriptLattes, a redução do trabalho por parte dos pesquisadores e dos profissionais de informação com o povoamento do repositório institucional se valendo do DOI, a partir de dados extraídos do Currículo Lattes para tal; e a aceleração do processo de povoamento do Repositório Institucional.

**Palavras-chave:** Repositório Institucional, Via Verde, Digital Object Identifier (DOI)

## Referências bibliográficas:

- GUÉDON, J.-C. (2004) – The “Green” and “Gold” roads to Open Access: The case for mixing and matching. *Serials Review*, Vol. 30, N° 4 (2004), p. 315-328
- MENA-CHALCO, J. P.; CESAR JUNIOR, R. M. (2009) – ScriptLattes: an open-source knowledge extraction system from the Lattes platform. *Journal of the Brazilian Computer Society*, Vol. 15, N° 4 (2009), p. 31-39
- SAYÃO, L. F. (2010) – Uma outra face dos metadados: informações para a gestão da preservação digital. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, Vol. 15, N° 30, p. 1-31