

# O MAPA DO TESOURO

Plataforma permite busca a dados de pesquisas realizadas em oito instituições de São Paulo

Rodrigo de Oliveira Andrade

Uma plataforma computacional lançada em dezembro permitirá que pesquisadores do Brasil e do exterior possam encontrar e acessar com mais facilidade dados científicos gerados em estudos desenvolvidos nas principais universidades públicas de São Paulo. A Rede de Repositórios de Dados Científicos do Estado de São Paulo é resultado de um esforço iniciado em julho de 2017 por oito instituições públicas de ensino e pesquisa paulistas: as três universidades estaduais, as três federais, o Instituto Tecnológico de Aeronáutica (ITA) e a Embrapa Informática Agropecuária.

“Trata-se de uma iniciativa inédita na América Latina em termos de complexidade e abrangência”, diz a cientista da computação Claudia Bauzer Medeiros, do Instituto de Computação da Universidade Estadual de Campinas (Unicamp) e coordenadora da iniciativa. “O sistema vai disponibilizar dados de pesquisas em todas as áreas do conhecimento hospedados em diferentes repositórios institucionais, construídos a partir das especificações tecnológicas e administrativas de cada instituição participante.”

A plataforma funciona como um sistema de busca, disponível no endereço

metabuscador.sc.usp.br. Com base em uma ou mais palavras-chave, ele percorre os repositórios de cada instituição participante. Em poucos segundos, reúne em uma lista uma relação de arquivos com dados de pesquisas associada aos termos procurados. Ao selecionar a informação de interesse, o usuário é redirecionado para o repositório da instituição que hospeda aquele conjunto de informações, podendo acessá-lo na íntegra – exceto nos casos de restrições éticas, de privacidade ou propriedade intelectual.

O objetivo é estimular o compartilhamento dessas informações e permitir sua reutilização em novas pesquisas ou na reprodução de experimentos. Para Medeiros, a rede deverá facilitar o acesso a esse tipo de informação, além de impulsionar parcerias, acelerar novas descobertas científicas e ampliar a visibilidade da produção acadêmica paulista. “Trata-se da chamada colaboração científica mediada por dados, que vem crescendo no mundo nos últimos tempos”, esclarece a cientista da computação.

Há alguns anos, as universidades públicas de São Paulo começaram a construir repositórios institucionais nos quais sua produção científica — artigos, teses e dissertações — pudesse ser consulta-

da. O da Universidade Estadual Paulista (Unesp), lançado em 2013, por exemplo, tem cópias de 39,8 mil artigos publicados em acesso aberto e 37 mil em acesso restrito. Já o da Unicamp reúne, desde 2015, mais de 155 mil documentos. Também a Universidade Federal de São Paulo (Unifesp) há quase cinco anos abastece seu repositório com a produção de seus pesquisadores. Seu acervo hoje é composto por 24,5 mil artigos, 11 mil dissertações de mestrado e 7 mil teses de doutorado (ver Pesquisa FAPESP nº 282). A Universidade de São Paulo (USP) também disponibiliza teses e dissertações desde 2001 – atualmente são mais de 88 mil títulos.

Mais recentemente, contudo, as instituições passaram a se preocupar também em fornecer seus dados de pesquisa, por meio da criação de repositórios de metadados, isto é, dados que informam a origem de outros dados. Os metadados oferecem uma descrição sobre as informações científicas geradas em determinada pesquisa, detalhando como foram produzidas, quando, onde e como podem ser reutilizadas e quem as gerou. “Essa movimentação foi estimulada em grande medida pela FAPESP”, destaca a cientista da computação Fátima Nunes,

professora da Escola de Artes, Ciências e Humanidades da USP, e uma das responsáveis pelo desenvolvimento da nova rede de repositórios de dados científicos.

**E**m outubro de 2017 a Fundação passou a exigir que os pesquisadores anexassem aos seus pedidos de financiamento de projetos um plano de gestão de dados científicos, desde a coleta até onde eles seriam disponibilizados. O objetivo era facilitar o trabalho de outros pesquisadores interessados em reproduzir os mesmos experimentos e racionalizar o uso de recursos públicos, potencializando o uso de informações já existentes. A demanda veio na esteira de iniciativas promovidas em países da Europa e nos Estados Unidos, e baseadas na ideia de que as pesquisas produzidas com recursos públicos devem ter seus resultados franqueados sem restrições, inclusive em relação aos dados coletados.

“A exigência da Fundação fez com que as instituições públicas de ensino e pesquisa de São Paulo trabalhassem juntas em políticas e diretrizes comuns de gerenciamento de informações de suas pesquisas”, esclarece Nunes. “Durante as reuniões que realizamos, chegou-se à conclusão de que cada instituição participante deveria criar e gerenciar seu próprio repositório de dados primários de pesquisa, e que eles deveriam ser alimentados pelos próprios pesquisadores.” Ao mesmo tempo, iniciaram o planejamento para desenvolver uma plataforma capaz de rastrear e compilar todas essas informações em um único lugar.

Coube à Superintendência de Tecnologia da Informação da USP o desafio de desenvolver o sistema que sustenta a rede de repositórios. A expectativa agora é a de que a plataforma ofereça aos gestores públicos de ciência e tecnologia uma visão ampla e detalhada da produção científica de São Paulo. “Espera-se também que ela se converta em uma ferramenta que possa ser usada pelas universidades na prestação de contas de suas atividades à sociedade”, comenta Medeiros.

Outro desafio que se apresenta com o lançamento da nova rede é o de incutir na rotina de trabalho dos pesquisadores a cultura do compartilhamento e reúso de dados científicos — os resultados das buscas na nova plataforma dependem da quantidade de informações disponibilizadas nos repositórios das instituições. O reaproveitamento de dados vem crescendo, mas ainda está longe de se consolidar no ambiente acadêmico (ver Pesquisa FAPESP nº 265).

A engenheira química Sônia Maria Malmonge, pró-reitora de Pesquisa da Universidade Federal do ABC (UFABC), uma das instituições que integra a rede de repositórios, declara-se confiante no potencial da plataforma em estimular o compartilhamento de dados. “Esperamos que a rede contribua para que os pesquisadores incorporem essa prática no seu dia a dia”, diz. A UFABC trabalha em uma política institucional de ciência aberta para incentivar seus pesquisadores nesse sentido, além de oferecer uma infraestrutura que viabilize o compartilhamento dos dados de suas pesquisas. ■

