

***CROSSREF EVENT DATA*: registro de atividades pós-publicação**

Edilson Damasio
Universidade Estadual de Maringá, Maringá, PR, Brasil
edamasio@uem.br

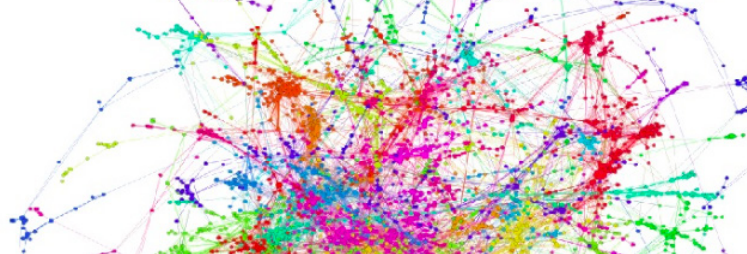
Rachael Lammey
Crossref, Oxford, Reino Unido
rlammey@crossref.org

1 INTRODUÇÃO

Editores, pesquisadores, financiadores, instituições e provedores de tecnologia estão interessados em entender melhor como a pesquisa acadêmica e seus resultados e dados são utilizados. O conteúdo acadêmico sempre foi discutido por estudiosos dentro da literatura formal usando as citações e referências, mas também é debatido fora da literatura formal e por outros meios além da comunidade acadêmica. Precisamos de maneiras de monitorar e distribuir essas informações valiosas. Assim, identificar se a publicação está gerando um impacto fora da literatura formal em diferentes plataformas de informação científica e suas relações com o crescimento das plataformas *on-line* de discussão, publicação e mídia social, é um dos grandes desafios dos editores/provedores/financiadores da ciência. Para isso, em 2017 foi iniciado um projeto de serviço de cruzamento entre links *Digital Object Identifiers* (DOIs), que está na versão Beta:

Uma iniciativa colaborativa do Crossref e DataCite¹, o *Event Data* atende a uma crescente necessidade da comunidade, ofe-

¹ Organização global sem fins lucrativos que fornece identificadores persistentes (DOIs) para dados de pesquisa, com objetivo de localizar, identificar e citar dados (DATA CITE, 2018).



recendo transparência em torno dos caminhos das interações de como as pesquisas acadêmicas ocorrem online, fora das plataformas do editor (CROSSREF, 2017a, tradução nossa).

Como exemplo, os DOIs de dados de pesquisa são registrados no *Data Cite* e outros locais de publicação com referências, tais como o *WordPress*, a *Wikipedia* e o *Crossref*, e também menções no *Twitter*, que geram dados bibliográficos e de mídia social depositados.

Quando surgem as interações, que são constituídas por uma série de “eventos” individuais registrados quando há um novo relacionamento entre o conteúdo registrado e algo na web, como um marcador/marcação, um comentário, um compartilhamento em redes sociais ou um link. O *Event Data* agrupa a relação desses dados e contabiliza dezenas de milhares de eventos ocorrendo todos os dias a uma taxa de aproximadamente um por segundo, e existe uma riqueza de informações que podem ser obtidas através da análise e interpretação desses dados (CROSSREF, 2017b).

A metodologia utilizada pelo *Event Data* consiste em registrar e armazenar essas atividades e torná-las disponíveis num modelo de registro de metadados bruto, para que qualquer interessado veja como o conteúdo já registrado no *Crossref* pode ser utilizado no mundo *on-line*. Os dados são transparentes, rastreáveis e seguem as recomendações da *National Information Standards Organization* (NISO) e do Código de Conduta da *Altmetrics Data Quality*, possibilitando o acesso a informações da proveniência e o contexto de cada “evento” junto ao serviço (CROSSREF, 2017b). Conforme o *Crossref*, as principais vantagens do serviço são: os dados são abertos, transferíveis e portáteis; as atualizações são regulares e os contribuintes confiáveis, com armazenamento seguro; o sistema fornece dados brutos gratuitamente, através de uma *Service Level Agreement* (SLA) (Acordo de Nível de Serviço) em velocidade favorável de disseminação; é uma iniciativa em comunidade, que fornece dados de uma fonte neutra e fidedigna.

A Figura 1 mostra a atividade ao vivo da coleta de dados no *Event Data*, realizada automaticamente em tempo real e que relaciona as fontes que incorpora; o tipo de “evento” fica no lado esquerdo (referências,



Twitter, comentário, atualização, entre outros), já a fonte fornecedora dos dados, que nessa imagem são DOIs do *Wikipedia*. fica no lado direito.

FIGURA 1 - COLETA EM TEMPO REAL DE DADOS NO *CROSSREF EVENT DATA*

Demo of Events flowing through Crossref Event Data



This is a demo designed to exhibit the stream of data that is flowing through Crossref Event Data at any given time. As is the nature with demos, it will sometimes go down. If the demo is down, it does not necessarily mean that Crossref Event data itself is down.

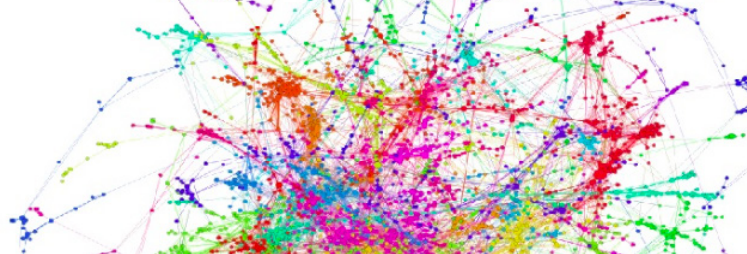
Live data from all sources.

wikipedia Enchytraeidae https://species.wikimedia.org/wiki/Enchytraeidae	references 05/02/2017	wikipedia https://doi.org/10.11646/zootaxa.4254.3.2
wikipedia Fossil https://de.wikipedia.org/wiki/Fossil	references 05/02/2017	wikipedia https://doi.org/10.1038/nature07671
wikipedia Enchytraeidae https://species.wikimedia.org/wiki/Enchytraeidae	references 05/02/2017	wikipedia https://doi.org/10.11646/zootaxa.3647.2.4
wikipedia Fossil https://de.wikipedia.org/wiki/Fossil	references 05/02/2017	wikipedia https://doi.org/10.1038/nature21377
wikipedia Enchytraeidae https://species.wikimedia.org/wiki/Enchytraeidae	references 05/02/2017	wikipedia https://doi.org/10.11646/zootaxa.3947.4.3

Fonte: Disponível em: <<https://live.eventdata.crossref.org/live.html>>. Acesso em: 05 maio 2018.

O *Event Data* e sua utilização pela comunidade científica tem o objetivo de explicar ao usuário como serão exploradas as suas métricas. O *Crossref* descreve alguns exemplos e possibilidades:

- Editores e plataformas de publicação podem analisar e interpretar os dados, utilizando os registros de eventos em análises métricas no seu gerenciamento comercial;
- É um recurso gerenciado de forma centralizada, aos terceiros (como prestadores de serviços em geral e altmétricos) que podem coletar dados em tempo real, para analisar, interpretar, enriquecer e relatar através de suas próprias ferramentas;



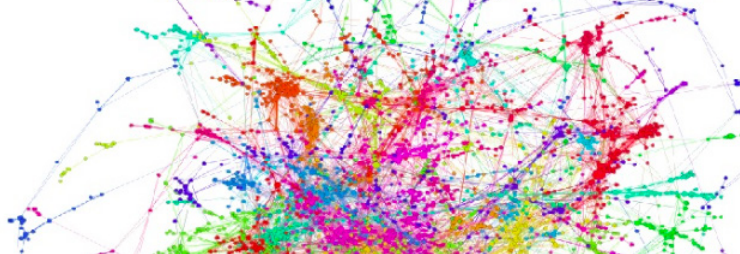
- Pesquisadores em bibliometria podem usar os dados brutos confiáveis e auditáveis em um único formato normalizado, com a certeza de que todos os dados correspondem à evidência e ao contexto da fonte;
- Editores de revistas podem utilizar o *Event Data* para muitas possibilidades, como: encontrar revisores com base em análise da rede de publicação; identificar novas áreas e aumentar o número de submissões; rastrear o alcance das submissões que vão publicar; atrair autores oferecendo dados do interesse dos leitores; acompanhar o alcance total da disseminação do artigo; obter melhor compreensão de como suas publicações se comparam com outras;
- Financiadores para identificar, isolar e rastrear a disseminação e o uso da pesquisa que financiaram fora da literatura acadêmica (CROSSREF, 2017a).

2 COMO O SERVIÇO É REALIZADO

2.1 COLETA DE DADOS NO *CROSSREF*

Uma vez registrado um conteúdo no *Crossref*, há diversas maneiras de acompanhá-lo. Com o DOI, pode-se descobrir quando um item foi compartilhado, salvo, referenciado, comentado ou *liked* (visto/apreciado) em ambientes variados como o *Twitter*, blogs e sites como o *Wikipedia* e *Research Blogging* (plataforma agregadora de blogs de pesquisadores). Com o *Event Data* também é possível capturar os links entre os autores, os conjunto de dados e outros itens com conteúdo na medida em que são adicionados aos metadados dos principais provedores do serviço, o *DataCite* e *Crossref* (CROSSREF EVENT DATA, 2018a).

De acordo com o *Crossref*, o serviço continuará se conectando a novas fontes de dados, para crescimento da atividade com o tempo. Assim, os “eventos” serão coletados e organizados a partir de fontes de dados, e alguns serão produzidos por meio de parcerias.



2.2 COLETA DE DADOS DOS PARCEIROS

Os dados do *Event Data* são fornecidos das fontes mencionadas na Figura 2. O *Crossref* os processa com base em uma metodologia (*schema*) e, em seguida, fornece os “eventos” e todas as informações do processamento dos dados para qualquer usuário ou pessoa através de uma própria Interface de Programação de Aplicativos (API), como um usuário final. O *Event Data* é adequado para utilização por máquinas/servidores em organizações que desejam criar os dados em sistemas próprios.

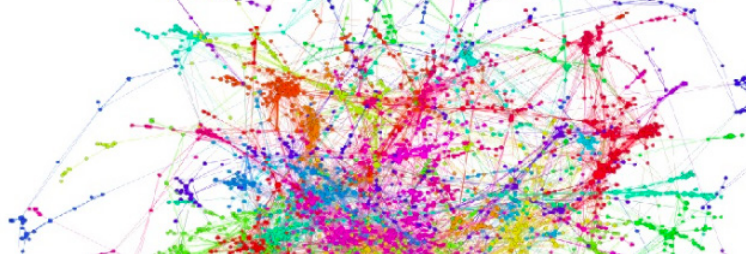
FIGURA 2 - FONTES DE DADOS COLETADAS PELO EVENT DATA (JAN. 2018)

Data source	Event type
Cambia Lens	Citations in Patents
Crossref Metadata	Links to DataCite registered content
DataCite Metadata	Links to Crossref registered content
Hypothes.is	Annotations in Hypothes.is
Newsfeed	Discussed in blogs and media
Reddit	Discussed on Reddit
Reddit Links	Discussed on sites linked to in subreddits
Stack Exchange Network	Discussed on StackExchange sites
Twitter	Mentions in tweets
Wikipedia	References on Wikipedia pages
Wordpress.com	Discussed on Wordpress.com sites

Fonte: Disponível em: <<https://www.crossref.org/services/event-data/>>. Acesso em: 05 maio 2018.

3 CONCLUSÃO

O serviço (abril de 2018) está em uma versão Beta e será lançado no primeiro semestre de 2018. Como uma metodologia de agregação de metadados abertos, o principal objetivo é identificar e ver como a



comunidade usa e integra esses metadados abertos nos fluxos de trabalho, por exemplo: o *Public Knowledge Project* (PKP) está trabalhando em uma implementação para o *Open Journal Systems* (OJS). Segundo o *Crossref*, o serviço precisa de comentários e sugestões das fontes que deveriam ser incluídas e outras possibilidades para o serviço. É necessário identificar após o lançamento quais serviços agregaram os dados *Event Data* e como eles estão sendo utilizados, pois na metodologia disponibilizada pelo *Crossref*, é através de um acordo de serviço. A ligação de diferentes identificadores persistentes para a coleta e distribuição de metadados via API, podem gerar muitas possibilidades de ferramentas métricas e altmétricas. É necessário que a comunidade examine a documentação disponibilizada no *Guia do usuário* e se informe dos termos para, em seguida, conectar-se à API do *Event Data* e começar a utilizá-la em outros desenvolvimentos.

AGRADECIMENTOS

A equipe que trabalha no *Event Data*, composta por Joe Wass, Madeline Watson, Jennifer Kemp (*Crossref*) e Martin Fenner (*DataCite*).

REFERÊNCIAS

CROSSREF. **Event Data**: open for you interpretation. 2017a. Disponível em: <<https://www.crossref.org/services/event-data/>>. Acesso em: 31 Jan. 2018.

_____. **Event Data terms of use**. 2017b. Disponível em: <<https://www.crossref.org/services/event-data/terms/>>. Acesso em: 31 Jan. 2018.

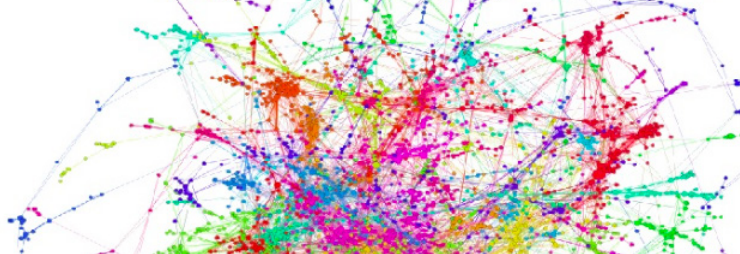
CROSSREF EVENT DATA. **Event Data user guide**. 2018a. Disponível em: <<http://www.eventdata.crossref.org/guide/index.html>>. Acesso em: 31 Jan. 2018.

_____. **Live feed of events being collected by Crossref Event Data**. 2018b. Disponível em: <<https://live.eventdata.crossref.org/live.html>>. Acesso em: 31 Jan. 2018.

DATA CITE. Disponível em: <<https://www.datacite.org/>>. Acesso em: 31 Jan. 2018.

6° EBBC

Rio de Janeiro
17 a 20 de julho



WASS, Joe. Bridging identifiers at PIDapalooza. **Crossref Blog**, 22 Jan. 2018. Disponível em: <<https://www.crossref.org/blog/bridging-identifiers-at-pidapalooza/>>. Acesso em: 12 Apr. 2018.

WATSON, Madeleine. Event Data: open for your interpretation. **Crossref Blog**, 25 Fev. 2016. Disponível em: <<https://www.crossref.org/blog/event-data-open-for-your-interpretation/>>. Acesso em: 31 Jan. 2018.