Metadata for data citation: what do I need to do?
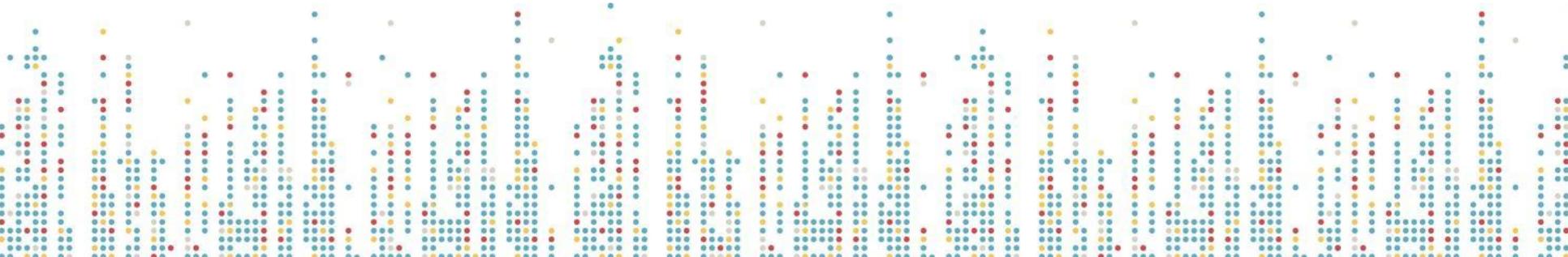
Citing data in the Crossref metadata

- Publishers regularly pass metadata and identifiers to Crossref when registering DOIs upon article publication
- The metadata includes bibilographic information about the content: title, publication dates, author details, funder information, ORCID iDs, license information, references and more.
- **Citations to data** should be added to the metadata publishers register with Crossref, using identifiers for the data (DOIs when possible).
- We'll then take the metadata and disseminate it in ways compatible with the Scholix framework, so that it can be used by those interested in these important links.

How does it all fit together?

Publish paper linking clearly to the data

Develop a data policy that includes data citation

Explain to authors how they should be citing data

Collect the data and pass it through production workflows

Send data to Crossref

Stuck for ideas? http://doi.org/10.1629/uksg.463

Basic

Advanced

Get journal buy-in

Use instructions for authors

Custom questions in manuscript submission systems

submission systems working on integrations with repositories

Copyeditors/typesetters/vendors

Available via REST API

Available via Event Data

SCHOLIX

Cousijn, H. *et al*. A data citation roadmap for scientific publishers. *Sci. Data*. 5:180259 https://doi.org/10.1038/sdata.2018.259 (2018).

Let's focus on this piece:
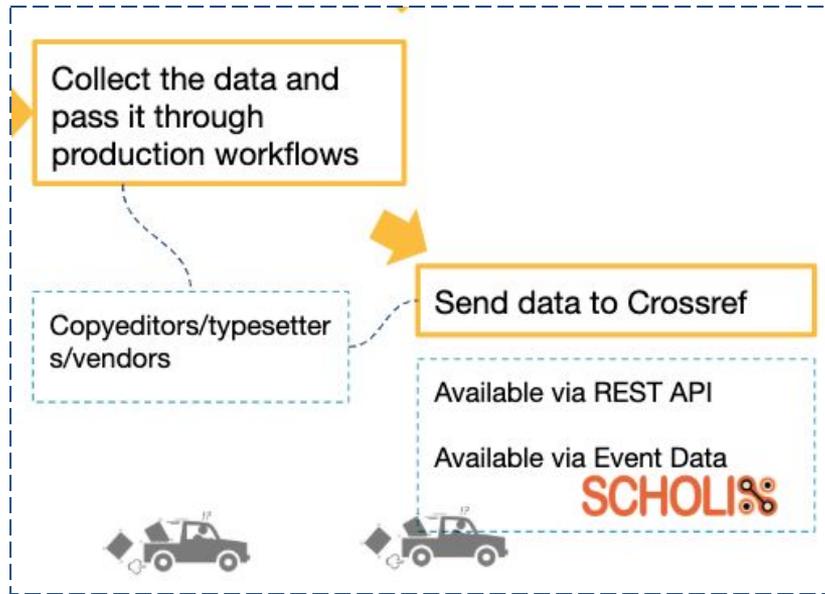
- Links to data needs to be sent to Crossref in one of a number of standard formats, ideally in reference lists
- That way we can recognize it as a data citation and make it available via the Event Data API

How do my data citations get into tools that use the Scholix framework from Crossref?

Crossref metadata

Crossref Event Data API

Subset of links to DataCite DOIs from Crossref's reference metadata

Event Data endpoint that provides this data in the Scholix forma

ScholeXplorer

SCHOLIX API
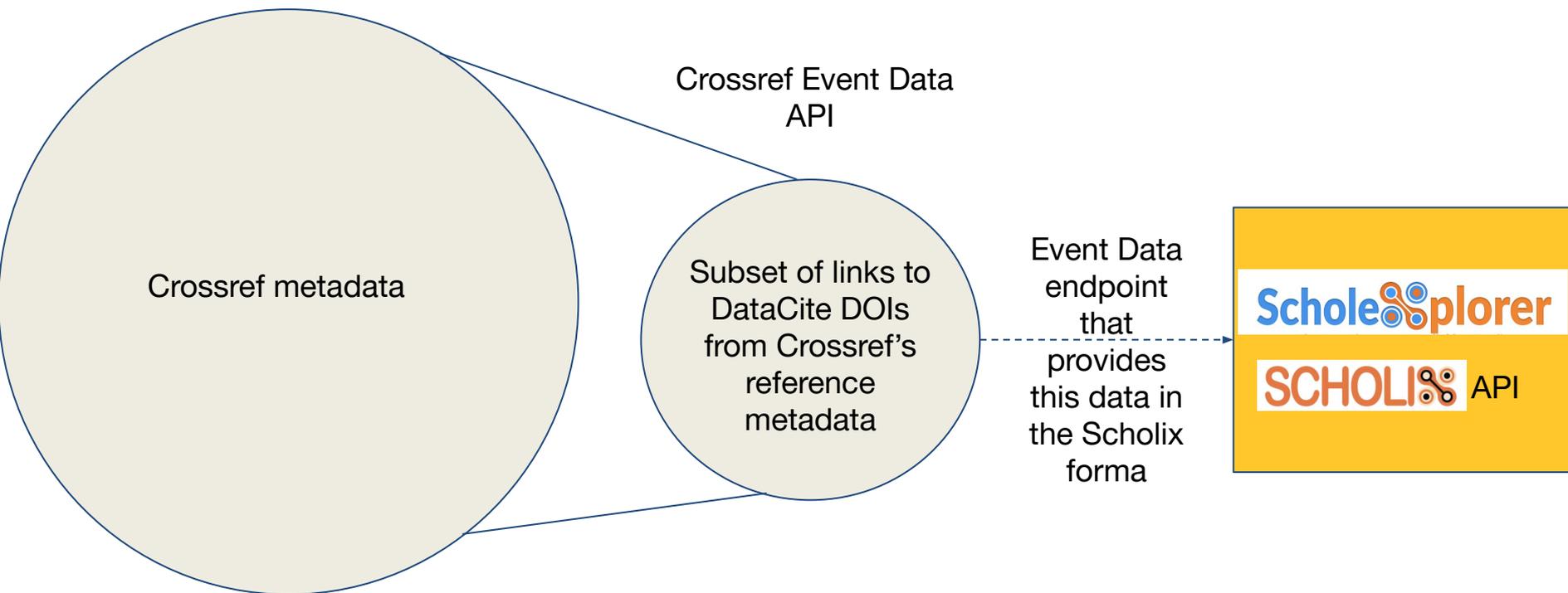
Extracting data citations into a separate service means they can easily be found.

Including data citations in Crossref metadata

1. References - include data citations in the references you register with Crossref. Can be as simple as just providing the DOI for the dataset:

```
<citation key="ref2">
  <doi>10.6084/m9.figshare.5981968</doi>
</citation>
```

Crossref recognizes that the DOI is for a DataCite DOI and pass it to Event Data.

Make sure your typesetters, copy editors and hosting platforms are briefed on how to work with these links and don't remove them.

Including data citations in Crossref metadata

2. Relations - include relationships between DOIs and other items in your Crossref metadata records

```
<rel:related_item>
  <rel:description>Sicard-2018-External-database-S1
  </rel:description>
  <rel:inter_work_relation identifier-type="doi"
  relationship-type="isSupplementedby">
   10.6084/m9.figshare.5981968
  </rel:inter_work_relation>
</rel:related_item>
```

Relations metadata is currently available via Crossref's REST API, and will be added into Event Data in future.

Future plans (next schema release)

Expanded citation markup with publication types:

```xml
<citation key="ref3" publication_type="data">
  <author>Morinha F</author>
  <cYear>2017</cYear>
  <institution>Dryad Digital Repository</institution>
  <title>Data from: Extreme genetic structure in a social bird species despite high dispersal capacity</title>
  <doi>10.5061/dryad.684v0</doi>
  <identifier type="accession">ABC123</identifier>
  <unstructured_citation>Morinha F, Dávila JA, Estela B, Cabral JA, Frías Ó, González JL, Travassos P, Carvalho D, Milá B, Blanco G (2017) Data from: Extreme genetic structure in a social bird species despite high dispersal capacity. Dryad Digital Repository. http://dx.doi.org/10.5061/dryad.684v0</unstructured_citation>
</citation>
```

Thank you!
Questions?

Useful links:
https://www.crossref.org/blog/data-citation-what-and-how-for-publishers/
support@crossref.org