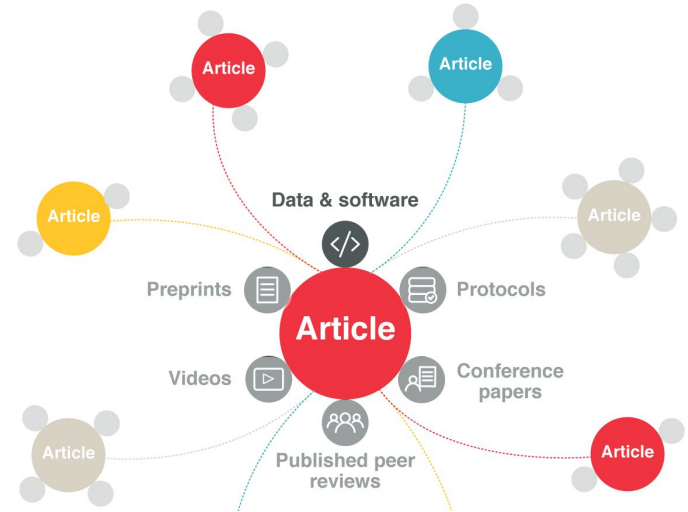


Crossref's role in the data citation workflow



Data citation for publishers

- Supports scholarship
 - Extends research
- Data cited consistently provides
 - Transparency
 - Context



“eLife is committed to ensuring researchers get credit for all their outputs, and data is a major component of this.” -- Melissa Harrison, eLife

How does it all fit together?

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

Publish paper linking clearly to the data

Stuck for ideas?
<http://doi.org/10.1629/uksq.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

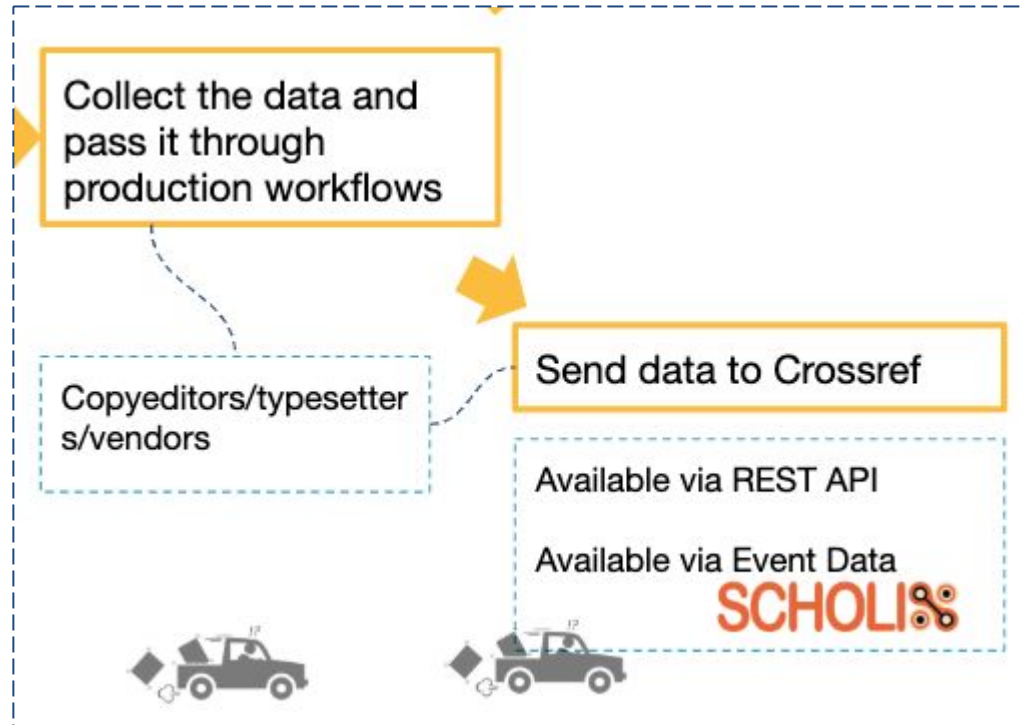
Send data to Crossref

Available via REST API

Available via Event Data

SCHOLIX





- Links to data needs to be sent to Crossref in one of a number of formats
- It's important for publishers to use these formats so that the data fits with the Scholix framework (and can therefore be used by Scholix).

Including data citations in Crossref metadata

1. References - include data citations in the citations you register with Crossref

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

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

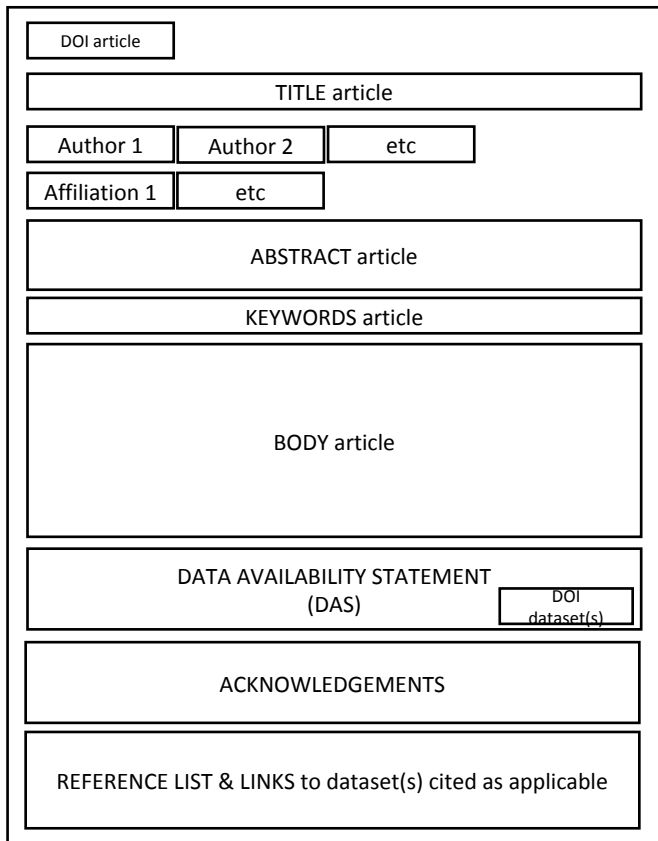
```
<rel:related_item>  
  <rel:description>Sicard-2018-External-database-S1  
</rel:description>  
  <rel:inter_work_relation identifier-type="doi"  
    relationship-type="isSupplementedby">  
    10.6084/m9.figshare.5981968  
  </rel:inter_work_relation>  
</rel:related_item>
```

Future plans

Expanded citation markup with publication types

```
<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>
```

How it looks like in the published article:



How it is covered in the underlying Crossref metadata:

```
<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>
```

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

Reference metadata

References



Data deposited in the Zenodo repository: URL: <https://zenodo.org/record/54705> with restricted access for reviewers (Palmer-Young *et al.*, 2016); at acceptance, data will be made freely available.

article

```
</citation>
▼<citation key="10.1111/jeb.13002-BIB0075|jeb13002-cit-0075">
  <author>Palmer-Young</author>
  <cYear>2016</cYear>
  ▼<article_title>
    Dataset: evolution of resistance to single and combined floral phytochemicals by a bumble bee parasite
  </article_title>
  <journal_title>Zenodo</journal_title>
  <doi>10.5281/zenodo.54705</doi>
</citation>
```

Crossref article metadata

```
"license": "https://doi.org/10.13003/CED-terms-of-use",
"obj_id": "https://doi.org/10.5281/zenodo.54705",
"source_token": "8676e950-8ac5-4074-8ac3-c0a18ada7e99",
"occurred_at": "2016-11-23T00:00:00Z",
"subj_id": "https://doi.org/10.1111/jeb.13002",
"id": "71f04fe8-5c15-4b5e-90ab-25e825ae820f",
"terms": "https://doi.org/10.13003/CED-terms-of-use",
"message_action": "create",
"source_id": "crossref",
"timestamp": "2017-05-19T13:30:17Z",
"relation_type_id": "references"
```

Crossref Event Data





Relationship metadata

SCIENTIFIC DATA 10110 01110 01110 01110 01110 01110 01110 01110 01110 01110

Data Descriptor | [Open Access](#) | Published: 22 October 2019

Satellite-based time-series of sea-surface temperature since 1981 for climate applications


Christopher J. Merchant , Owen Embury, Claire E. Bulgin, Thomas Block, Gary K. Corlett, Emma Fiedler, Simon A. Good, Jonathan Mittaz, Nick A. Rayner, David Berry, Steinar Eastwood, Michael Taylor,

 CEDA Archive

[Search Catalogue](#) [Get Data](#) [Help](#) [Tools](#) [Dep](#)

This website uses cookies. By continuing to use this website you are agreeing to our use of cookies.

Dataset

 European Space Agency
Sea Surface Temperature
Climate Change Initiative (ESA CCI SST) Logo

ESA Sea Surface Temperature Climate Change Initiative (SST_cci): Along-Track Scanning Radiometer (ATSR) Level 2 Preprocessed (L2P) Climate Data Record.

```
</CrossMark>
▼<rel:program xmlns:rel="http://www.crossref.org/relations.xsd" name="relations">
  ▼<rel:related_item>
    ▼<rel:description>
      Refractive indices (500–3500 cm-1) and emissivity (600–3350 cm-1) of pure water and seawater
    </rel:description>
    <rel:inter_work_relation identifier-type="doi" relationship-type="references">10.7488/ds/162</rel:inter_work_relation>
  </rel:related_item>
  ▼<rel:related_item>
    ▼<rel:description>
      ESA Sea Surface Temperature Climate Change Initiative (SST_cci): Along-Track Scanning Radiometer (ATSR) Level 2 Preprocessed (L2P) Climate Data Record, version 2.1
    </rel:description>
    <rel:inter_work_relation identifier-type="doi" relationship-type="references">10.5285/916b93aaf1474ce793171a33ca4c5026</rel:inter_work_relation>
  </rel:related_item>
  ▼<rel:related_item>
    ▼<rel:description>
      ESA Sea Surface Temperature Climate Change Initiative (SST_cci): Advanced Very High Resolution Radiometer (AVHRR) Level 2 Preprocessed (L2P) Climate Data Record, version 2.1
    </rel:description>
    <rel:inter_work_relation identifier-type="doi" relationship-type="references">10.5285/373638ed9c434e78b521cbe01ace5ef7</rel:inter_work_relation>
  </rel:related_item>
```

References

- Good option for publishers who make their reference metadata openly available via Crossref
- Can be a better fit with their existing workflows
- Data Citations made openly available via Crossref's APIs
- Citations with DataCite DOI are sent to Event Data

Relations

- At the moment, these can give more context about the data/article relationship (can specify 'references' or 'isSupplementedBy' as relation type)
- Good option for publishers who don't make their reference information visible via Crossref
- Relations are available via metadata APIs
- Future: will be events in Event Data



The case of the missing metadata

The screenshot shows a journal article page with a navigation menu on the left. The 'Supplemental material' tab is selected. The main content area features a 'Related Research Data' section with two entries, each containing a snippet of text and a 'Source: Figshare' link. At the bottom of this section, it says 'Linking provided by ScholarSplorer'.

Supplemental data

Supplemental data for this article can be accessed at the Open Science Framework, at <https://osf.io/nft78/>.

Data availability statement

The data described in this article are openly available in the Open Science Framework at <https://osf.io/nft78/>

No DOI in the metadata

The browser window shows the URL `api.crossref.org/works/10.1080/00224545.2018.1498317/transform/application/vnd.crossref.unixsd+xml`. The address bar also shows a link to `https://osf.io/nft78/`. The XML content is as follows:

```
mark>
  <crossmark_version>1</crossmark_version>
  <crossmark_policy>10.1080/tandf_crossmark_01</crossmark_policy>
  <crossmark_domains>
    <ossmark_domain>
      <domain>www.tandfonline.com</domain>
      <rossmark_domain>
        <ssmark_domains>
          <ssmark_domain_exclusive>true</crossmark_domain_exclusive>
        </ssmark_domains>
      </rossmark_domain>
    </ossmark_domain>
  </crossmark_domains>
  <om_metadata>
    <sertion name="peerreview_statement" label="Peer Review Statement" order="1">
      he publishing and review policy for this title is described in its Aims & Scope.
    </sertion>
    <sertion name="aims_and_scope_url" label="Aim & Scope" order="2" href="http://www.tandfonline.com/action/journalInformation?show=aimsScope&journalCode=vsoc20">
      http://www.tandfonline.com/action/journalInformation?show=aimsScope&journalCode=vsoc20
    </sertion>
  </om_metadata>
</mark>
```



The case of the missing metadata

observations.

Acknowledgments

We thank two anonymous reviewers for helpful comments. This research was supported by the Integrated Research Program for Advancing Climate Models (TOUGOU program) and the FLAGSHIP 2020 Project from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. The simulations in this study were performed using the K computer (hp120279, hp130010, and hp140219) and the Earth Simulator. The data used to generate the figures can be found at [https://figshare.com/articles/data for Noda et al JAMES/7581824](https://figshare.com/articles/data_for_Noda_et_al_JAMES/7581824) and are available upon request by contacting the corresponding author.

No use of DOI, no metadata :(

```
- Calendar SugarCRM - Comm... CrossRef Administr... crossref.org The Secret Life of... figshare| 0/0 x sf.org/wo... Other Bookm
appear to have any style information associated with it. The document tree is shown below.
ns="http://www.crossref.org/qrschema/3.0" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" version="3.0" xsi:schemaLocation="http://www.crossref.org/qrschema/3.0
org/schemas/crossref_query_output3.0.xsd">
one</doi_batch_id>
resolved">
urnal_article">10.1029/2019MS001658</doi>
="publisher-name" type="string">American Geophysical Union (AGU)</crm-item>
="prefix-name" type="string">American Geophysical Union</crm-item>
="member-id" type="number">13</crm-item>
="citation-id" type="number">109751310</crm-item>
="journal-id" type="number">85606</crm-item>
="deposit-timestamp" type="number">2019102415232600452</crm-item>
="owner-prefix" type="string">10.1029</crm-item>
="last-update" type="date">2019-10-24T22:23:58Z</crm-item>
="created" type="date">2019-09-12T14:00:21Z</crm-item>
="citedby-count" type="number">0</crm-item>
lns="http://www.crossref.org/xschema/1.1" xsi:schemaLocation="http://www.crossref.org/xschema/1.1 http://doi.crossref.org/schemas/unixref1.1.xsd">
_metadata language="en">
_title>Journal of Advances in Modeling Earth Systems</full_title>
_title>J. Adv. Model. Earth Syst.</abbrev_title>
edia_type="print">1942-2466</lsan>
edia_type="electronic">1942-2466</lsan>
l_metadata>
ssue>
```



The case of the missing metadata

71. Koch M, Quast B, Bartolomaeus T. Coeloms and nephridia in annelids and arthropods. In: Wägele W, Bartolomaeus T, editors. Deep Metazoan Phylogeny: The Backbone of the Tree of Life. Berlin: De Gruyter; 2014. p. 173–284.
72. Andrikou C, Thiel D, Ruiz-Santesteban JA, Hejnol A. Active mode of excretion across digestive tissues predates the origin of excretory organs. 2019. Dryad Digital Repository. <https://doi.org/10.5061/dryad.bq068jr>.
[View Article](#) • [Google Scholar](#)
73. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7. pmid:15034147

Reference to the data in
the reference list, no DOI
in the Crossref metadata

PLOS is definitely one of those publishers where it is mercifully easy to find the data in multiple places - it's amazing how and where the information can be hidden. In the example you sent, I can see that everything is there in the XML we have, bar the DataCite DOI:

```
<citation key="ref72">  
<journal_title>Dryad Digital Repository</journal_title>  
<author>C Andrikou</author>  
<cYear>2019</cYear>  
<article_title>  
Active mode of excretion across digestive tissues predates the origin  
of excretory organs  
</article_title>
```

Event Data Service is crucial for Data Citation DataCite FIND, ACCESS, AND REUSE DATA

Before Event Data

- Data citations needed to be found in DOI metadata, separate for Crossref and DataCite.
- Doesn't scale well, limiting adoption of data citation.

After Event Data

- Data citations are extracted into a separate service and can easily be found.
- Scales well, will lead to adoption of data citation.

European high-skilled mobility data and Scientific publication & collaboration data

Petersen, Alexander, UC Merced,  <https://orcid.org/0000-0002-0955-3483>

Publication date: March 4, 2017

Publisher: UC Merced

<https://doi.org/10.6071/M3RP49>

Citation

Petersen, Alexander (2017), European high-skilled mobility data and Scientific publication & collaboration data, UC Merced Dash, Collection, <https://doi.org/10.6071/M3RP49>

Abstract

Raw data files corresponding to longitudinal country-level observations for two types of variable: (i) publication, citation, and international collaboration data from SCImago Journal & Country Rank (www.scimagojr.com/) and (ii) high-skilled mobility data, corresponding to headcounts by source and destination country, from the "Professionals moving abroad (Establishment)" data hosted by the European Union's "The EU Single Market Regulated professionals database".

European high-skilled mobility data and Scientific publication & collaboration data

 Download the dataset ~ 3 MB

 Download Data Publication (PDF)

Petersen, Alexander M.; Puliga, Michelangelo (2017), High-skilled labour mobility in Europe before and after the 2004 enlargement, The Royal Society, Article-journal, <https://doi.org/10.1098/rsif.2017.0030>

Publicat
Publishe
https://c

Works Referencing This Dataset

Petersen, Alexander M.; Puliga, Michelangelo (2017), High-skilled labour mobility in Europe before and after the 2004 enlargement, The Royal Society, Article-journal, <https://doi.org/10.1098/rsif.2017.0030>

Doria Arrieta, Omar A.; Pammolli, Fabio; Petersen, Alexander M. (2017), Quantifying the negative impact of brain drain on the integration of European science, American Association for the Advancement of Science (AAAS), Article-journal, <https://doi.org/10.1126/sciadv.1602232>

Citatio
Petersen
Merced

Abstract

Raw data files corresponding to longitudinal country-level observations for two types of variable: (i) publication, citation, and international collaboration data from SCImago Journal & Country Rank (www.scimagojr.com/) and (ii) high-skilled mobility data, corresponding to headcounts by source and destination country, from the "Professionals moving abroad (Establishment)" data hosted by the European Union's "The EU Single Market Regulated professionals database".

October 25, 2017

Metrics

 66 views

 4 downloads

 [2 citations](#)

Type your search!

Q Search

Search by persistent Identifier

Q Search PID



Talking to publishers

- Keen to do this and understand it's important
- Building awareness and encouraging adoption by their journals
- Honing workflows to make sure the information gets all the way to Crossref
- Making relationships available via Event Data will help those registering data/article links in that way



Thank you!
Questions?

Links to publisher content from the DataCite metadata

```
"status": "ok",
"message-type": "event-list",
"message": {
  "next-cursor": "f941cb7f-dcbf-460e-9aed-cb9c749491fe",
  "total-results": 31343,
  "items-per-page": 10,
  "events": [
    {
      "license": "https://creativecommons.org/publicdomain/zero/1.0/",
      "obj_id": "https://doi.org/10.1017/s0954102003001640",
      "occurred_at": "2005-02-13T03:25:02Z",
      "subj_id": "https://doi.org/10.1594/pangaea.126301",
      "id": "8c05cdc9-2008-4973-b0f4-18666f7a288d",
      "terms": "https://doi.org/10.13003/CED-terms-of-use",
      "message_action": "create",
      "source_id": "datacite",
      "timestamp": "2017-03-11T11:38:17Z",
      "relation_type_id": "references"
    },
    {
      "license": "https://creativecommons.org/publicdomain/zero/1.0/",
      "obj_id": "https://doi.org/10.1017/s0954102003001640",
      "occurred_at": "2005-02-13T03:25:06Z",
      "subj_id": "https://doi.org/10.1594/pangaea.126300",
      "id": "0ccd121c-e464-4c75-890c-d0c89303b72e",
      "terms": "https://doi.org/10.13003/CED-terms-of-use",
      "message_action": "create",
      "source_id": "datacite",
      "timestamp": "2017-03-11T11:38:17Z",
      "relation_type_id": "references"
    }
  ]
}
```

```
"status": "ok",
"message-type": "event-list",
"message": {
  "next-cursor": "6cddc78c-f8c8-42bd-bc72-2ed5cdb3257f",
  "total-results": 794564,
  "items-per-page": 10,
  "events": [
    {
      "license": "https://creativecommons.org/publicdomain/zero/1.0/",
      "obj_id": "https://doi.org/10.1016/j.epsl.2012.08.012",
      "occurred_at": "2017-03-10T03:09:34Z",
      "subj_id": "https://doi.org/10.1594/pangaea.871708",
      "id": "b917c0cd-ad78-453f-b08d-5b486a7df41",
      "terms": "https://doi.org/10.13003/CED-terms-of-use",
      "message_action": "create",
      "source_id": "datacite",
      "timestamp": "2017-03-11T11:36:00Z",
      "relation_type_id": "is_supplement_to"
    },
    {
      "license": "https://creativecommons.org/publicdomain/zero/1.0/",
      "obj_id": "https://doi.org/10.1016/j.jembe.2015.12.001",
      "occurred_at": "2017-03-10T03:09:13Z",
      "subj_id": "https://doi.org/10.1594/pangaea.871938",
      "id": "52f27749-ceed-4ae9-a8c9-397f0cbe7f8c",
      "terms": "https://doi.org/10.13003/CED-terms-of-use",
      "message_action": "create",
      "source_id": "datacite",
      "timestamp": "2017-03-11T11:36:00Z",
      "relation_type_id": "is_supplement_to"
    }
  ]
}
```



PANGAEA.

Data Publisher for Earth & Environmental Science

Citation:

Hillenbrand, Claus-Dieter; Ehrmann, Werner (2003): Sedimentology of core PS1812-6. *PANGAEA*, doi <https://doi.org/10.1594/PANGAEA.126301>,

In supplement to: Hillenbrand, C-D; Ehrmann, W (2003): Palaeoenvironmental Implications from Tertiary Sediments of Kainan Maru Seamount and Northern Gunnerus Ridge. *Antarctic Science*, **15**, 522-536, doi <https://doi.org/10.1017/S0954102003001640>

Always quote above citation when using data! You can download the citation

RIS Citation

BisTeX Citation

Copy Citation

Facebook

Twitter

Show

Antarctic Science

Search An

Article

Metrics

Volume 15, Issue 4 December 2003, pp. 522-536

Cited by 6

[Get access](#)

Palaeoenvironmental implications of Tertiary sediments from Kainan Maru Seamount and northern Gunnerus Ridge

CLAUS-DIETER HILLENBRAND ^(a1) and WERNER EH RMANN ^(a2)

DOI: <https://doi.org/10.1017/S0954102003001640> Published online by Cambridge University Press: 02 December 2003

Abstract

Sedimentary sequences spanning early Oligocene and Neogene time intervals were recovered with piston and gravity cores along erosional structures at northern Gunnerus Ridge and Kainan Maru Seamount in the southernmost Indian Ocean. Results of sedimentological investigations help to reconstruct the Cenozoic palaeoenvironment. Main emphasis was placed on grain size and clay mineral data. The clay mineral assemblages are dominated by illite and smectite. Chlorite and kaolinite occur in trace amounts. Whereas illite has a distinct source on the East Antarctic craton, smectite is of somewhat speculative origin, but probably is derived from erosion of Cenozoic or older shelf sediments. The presence of terrigenous sand indicates that ice-rafting was active throughout the time represented by

Event Data APIs

1

Crossref Event Data Query API

Query Event Data using a large number of parameters.

<https://www.eventdata.crossref.org/guide/>

2

DataCite REST API

Query Event Data that include a DataCite DOI. Integrated into the DataCite REST API, includes basic DOI metadata, and aggregations.

<https://support.datacite.org/docs/eventdata-guide>

3

Scholix-compatible API

Data citations in Event Data using the Scholix metadata standard.

Useful links and contacts:

- Crossref Data & Software Citation Deposit Guide for Publishers:
<https://support.crossref.org/hc/en-us/articles/215787303-Crossref-Data-Software-Citation-Deposit-Guide-for-Publishers>
- Crossref Event Data: <https://www.crossref.org/services/event-data/>
- A Data Citation Roadmap for Scientific Publishers:
<https://doi.org/10.1038/sdata.2018.259>
- A Data Citation Roadmap for Scholarly Data Repositories
<https://doi.org/10.1101/097196>
- Make Data Count webinar:
<https://www.youtube.com/watch?v=Lkysz0Mc7fo>
- support@crossref.org/support@datacite.org