

Impulsvortrag 4: ORCID-Integration in Forschungsdatenrepositorien am Beispiel von PANGAEA

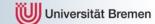
Uwe Schindler

PANGAEA (MARUM, University of Bremen)







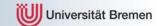


My Background

- Member of the PANGAEA team @ MARUM, University of Bremen.
- Studied **physics** long time ago.
- Responsible for metadata processing and search engine of PANGAEA (Elasticsearch).
- Long time Open Source software contributor; member of **Apache Software Foundation**: Apache Lucene, Apache Solr, Apache TIKA, Apache POI,... also **Elasticsearch /Opensearch**.







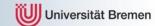
www.pangaea.de

About PANGAEA











Founded: 1993

Hosted by:

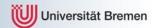




- 2001: Accreditation by the "International Council for Science" (ICSU) as "Publisher for Earth & Environmental Science" (ICSU WDS World Data Center)
- 2007: Accredited by the "World Meteorological Organisation" (WMO) as "World Radiation Monitoring Center" (WRMC) (since 2007)













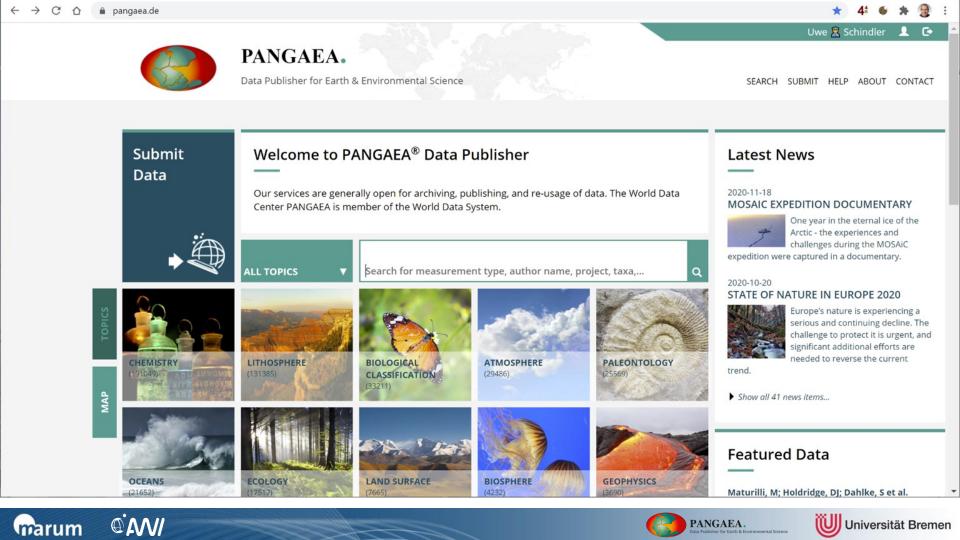












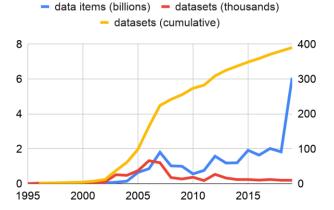


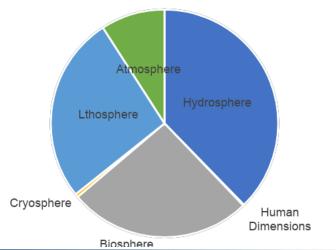
- Integral part of science: More than 250 European to international projects since 1995 (https://www.pangaea.de/projects)
- highly heterogeneous & dynamic
- multidisciplinary

Number of data sets ~402.000 Number of data items >19 Billion Data volume <3 PB Increase ~4% per year

>15.000 registered users 12 – 15.000 unique sessions / month

growth













PANGAEA – Data/Metadata Dissemination













EARTH OBSERVATIONS

GROUP ON















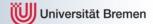












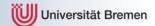
History

ORCID support in PANGAEA









July 2013

"Code Sprint" Workshop at GFZ Potsdam on July 1st, 2013 (DFG projects Komfor, Radieschen, re3data)

- Brainstorming (with **Martin Fenner**, PLoS): How to "claim" ORCID iDs *automatically* for PANGAEA's thousands of datasets
- First implementation went live on same day!







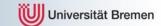


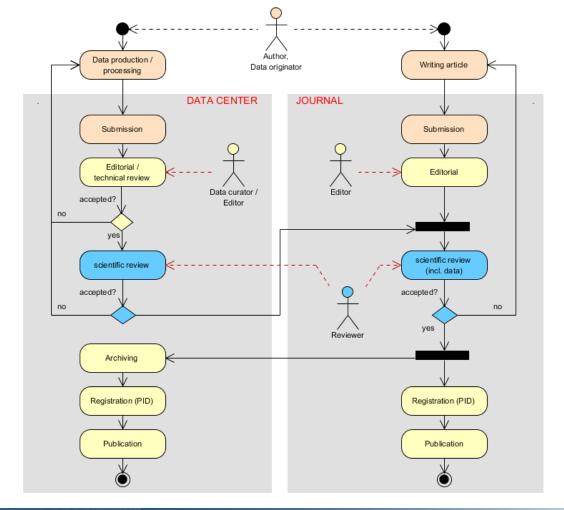
How?

- PANGAEA datasets are often linked to articles
- Journals take care of ORCID iD

















PANGAEA.

Data Publisher for Earth & Environmental Science

Citation:

Klages, Johann Philipp; Salzmann, Ulrich; Bickert, Torsten; Hillenbrand, Claus-Dieter; Gohl, Karsten; Kuhn, Gerhard; Bohaty, Steven M; Titschack, Jürgen; Müller, Juliane; Frederichs, Thomas; Bauersachs, Thorsten; Ehrmann, Werner; van de Flierdt, Tina; Simões Pereira, Patric; Larter, Robert D; Lohmann, Gerrit; Niezgodzki, Igor; Uenzelmann-Neben, Gabriele; Zundel, Maximilian; Spiegel, Cornelia; Mark, Chris; Chew, David M; Francis, Jane E; Nehrke, Gernot; Schwarz, Florian; Smith, James A; Freudenthal, Tim; Esper, Oliver; Pälike, Heiko; Ronge, Thomas A; Dziadek, Ricarda; Expedition P5104 Scientists (2019): Sedimentological, palynological, geochemical, palaeomagnetic, and geochronological investigations of cores 9R and 10R from MARUM-MeBo70 Site PS104_20-2. PANGAEA,

https://doi.org/10.1594/PANGAEA.906092

Always quote citation above when using data! You can download the citation in several formats below.

RIS Citation

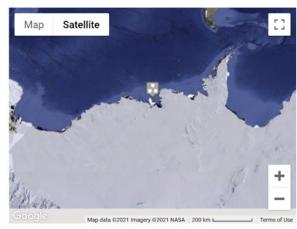
BIBTEX Citation 🙎 Copy Citation

C Facebook C Twitter

Show Map Google Earth







Abstract:

Here we present sedimentological, micropalaeontological, inorganic and organic geochemical, palaeomagnetic, geochronological as well as X-ray computed tomography (CT) data obtained from successional drill cores 9R and 10R at MARUM-MeBo70 Site PS104_20-2 (73.57°S, 107.09°W; 946 m water depth). The site is located within the Pine Island cross-shelf trough in the Amundsen Sea Embayment (ASE), West Antarctica (Fig. 1a) and was drilled during RV Polarstern Expedition PS104 in 2017. We further provide the palaeoclimate modeling analytical scripts related to this work.

Keyword(s):

climate Q; Late Cretaceous Q; palaeoenvironment Q; West Antarctica Q

Supplement to:

Klages, Johann Philipp; Salzmann, Ulrich; Bickert, Torsten; Hillenbrand, Claus-Dieter; Gohl, Karsten; Kuhn, Gerhard; Bohaty, Steven M; Titschack, Jürgen; Müller, Juliane; Frederichs, Thomas; Bauersachs, Thorsten; Ehrmann, Werner; van de Flierdt, Tina; Simões Pereira, Patric; Larter, Robert D; Lohmann, Gerrit; Niezgodzki, Igor; Uenzelmann-Neben, Gabriele; Zundel, Maximilian; Spiegel, Cornelia; Mark, Chris; Chew, David M; Francis, Jane E; Nehrke, Gernot; Schwarz, Florian; Smith, James A; Freudenthal, Tim; Esper, Oliver; Pälike, Heiko; Ronge, Thomas A; Dziadek, Ricarda; Expedition PS104 Scientists (2020): Temperate rainforests near the South Pole during peak Cretaceous warmth. Nature, 580, 81-86, https://doi.org/10.1038/s41586-020-2148-5 Q









How?

PANGAEA uses ORCID search (Apache Solr) with following query parameters:

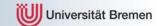
- All known article DOIs of author (recent first)
- Last name, first name with wildcard (to disambiguate)

Accepts only ONE search result!









Isn't this against ORCID terms?

- We have a full record of author (all datasets, all related articles, his coworkers,...)
- We know the article and its authors
- Article publisher / user has claimed and validated ORCID iD



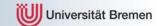


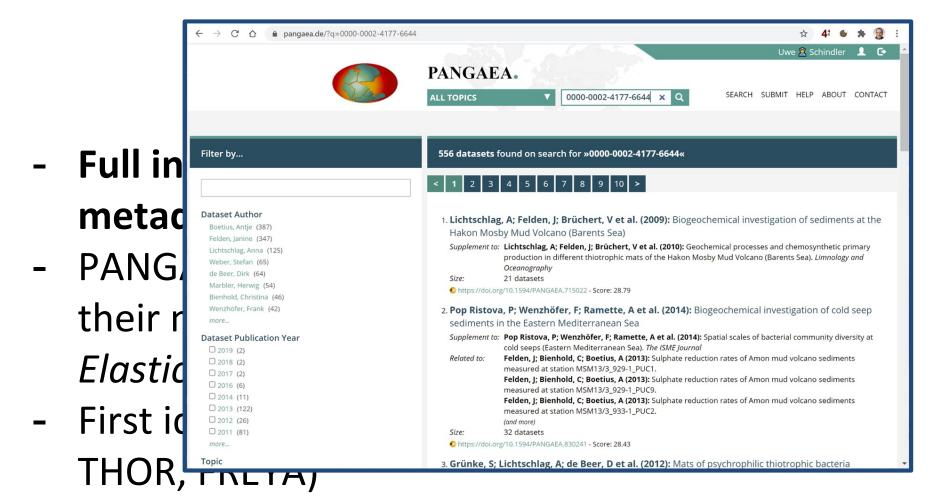
September 2013

- Full integration of ORCID iD into PANGAEA's metadata schema, transfer to DataCite
- PANGAEA allows to **search for ORCID iD** in their main search engine (Apache Lucene / Elasticsearch)
- First ideas of "PID Graph" (later EU projects: THOR, FREYA)

















September 2013

- Full integration of ORCID iD into PANGAEA's metadata schema, transfer to DataCite
- PANGAEA allows to **search for ORCID iD** in their main search engine (Apache Lucene / Elasticsearch)
- First ideas of "PID Graph" (later EU projects: THOR, FREYA)







April 2016: EU Project THOR

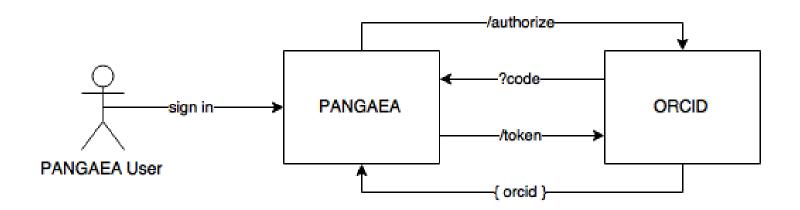
- Login and user registration possible via ORCID OAuth
- Going live: September 2016 with PANGAEA's new web page







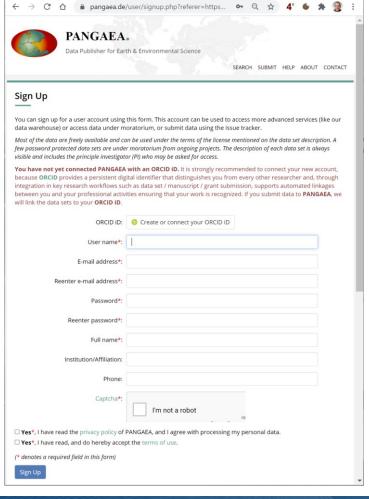
ORCID OAuth

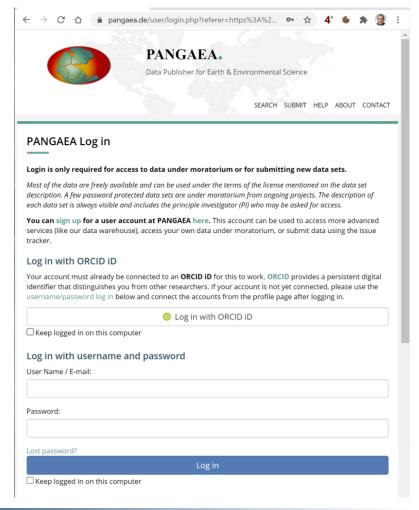








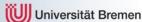


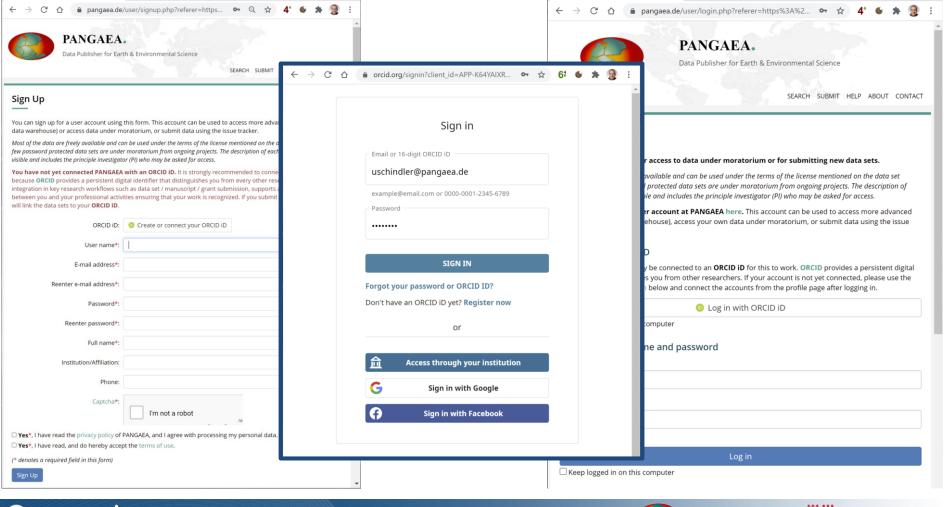










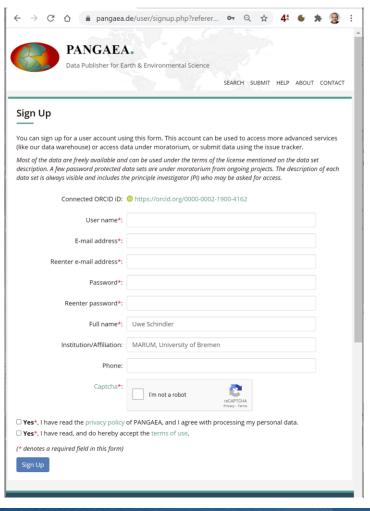


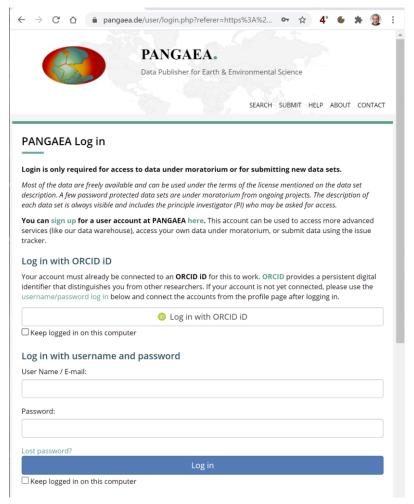




















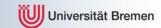


Live demo









ccessible nteroperable

Image credit: FAIR data principles by SangyaPundir at <u>Wikimedia Commons</u>











indable Accessible

ccessible nteroperable



****eusable



Image credit: FAIR data principles by SangyaPundir at <u>Wikimedia Commons</u>







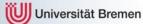


```
-bash
                                                                                                                                                                 X
Uwe Schindler@VEGA:~ > curl -D- -LH'Accept: application/ld+json' https://doi.pangaea.de/10.1594/PANGAEA.914629
HTTP/2 200
server: nainx/1.21.0
date: Tue, 01 Jun 2021 15:42:23 GMT
content-type: application/ld+json; charset=UTF-8
set-cookie: pansessid=7fc48273de6765617fb794fb81479a4e; Path=/; Domain=.pangaea.de; Secure; Httponly
vary: Cookie, X-PanLoginID, Authorization
cache-control: public
x-cid: 7fc48273de6765617fb794fb81479a4e
varv: Accept
|link: <https://doi.org/10.1594/PANGAEA.914629>;rel="cite-as", <https://doi.pangaea.de/10.1594/PANGAEA.914629>;rel="describes", <https://
9?format=zip>;rel="item";type="application/zip", <https://orcid.org/0000-0001-8646-2334>;rel="author", <https://orcid.org/0000-0002-6952
org/0000-0001-6053-2330>;rel="author", <https://orcid.org/0000-0002-4453-9564>;rel="author"
x-powered-by: Jetty(9.4.36.v20210114)
strict-transport-security: max-age=31536000
x-ua-compatible: IE=Edge
x-content-type-options: nosniff
x-frame-options: SAMEORIGIN
{"@context":"http://schema.org/","@id":"https://doi.org/10.1594/PANGAEA.914629","@type":"Dataset","identifier":"https://doi.org/10.1594/
angaea.de/10.1594/PANGAEA.914629","creator":[{"@id":"https://orcid.org/0000-0001-8646-2334","@type":"Person","name":"Walter Geibert","fa
ter","identifier":"https://orcid.org/0000-0001-8646-2334","url":"http://www.awi.de/ueber-uns/organisation/mitarbeiter/walter-geibert.htm
{"@id":"https://orcid.org/0000-0002-6952-2494","@type":"Person","name":"Jens Matthiessen","familyName":"Matthiessen","givenName":"Jens",
-0002-6952-2494", "url": "http://www.awi.de/en/about-us/organisation/staff/jens-matthiessen.html", "email": "jens.matthiessen@awi.de"},{"@id
2330","@type":"Person","name":"Ingrid Stimac","familyName":"Stimac","givenName":"Ingrid","identifier":"https://orcid.org/0000-0001-6053-
ta E wollenburg","familyName":"Wollenburg","givenName":"Jutta E","url":"http://www.awi.de/en/about-us/organisation/staff/jutta-wollenbur
i.de"},{"@id":"https://orcid.org/0000-0002-4453-9564","@type":"Person","name":"Ruediger Stein","familyName":"Stein","givenName":"Ruedige
000-0002-4453-9564","url":"http://www.awi.de/en/about-us/organisation/staff/ruediger-stein.html","email":"ruediger.stein@awi.de"}],"name
al data together with microfossil data and their stable isotopes","publisher":{"@type":"Organization","name":"PANGAEA","disambiguatingDe & Environmental Science","url":"https://www.pangaea.de/"},"includedInDataCatalog":{"@type":"DataCatalog","name":"PANGAEA","disambiguatiarth & Environmental Science","url":"https://www.pangaea.de/"},"datePublished":"2020-04-09","description":"This dataset comprises previo
omposition of sediment cores from the Arctic Ocean. New data are supplied for cores PS51/038-4, PS72/396-5 and PS72/396-3: For PS51/038-
 230Th excess, 238U, 234U/238U together with bulk Ca, Mn and S concentrations. For PS72/396-3, we provide data on planktonic foraminifer
composition, allochtonous forminifera fractions and the abundance of Cibicides lobatulus subsp. grossa. Thorium and uranium isotopes wer
(Element2) via isotope dilution; elemental concentrations on an ICP-OES iCAP; foraminfera abundances after freeze-drying and sieving of
le isotopés on a Finnigan MAT 251 isotope ratio gas mass spectrometer directly coupled to a Kiel I automatic carbonate preparation devic te. 230Thex was calculated by subtracting the activity of 2340 from 230Th. [...]","abstract":"This dataset comprises previously published
```









```
-bash
                                                                                                                                     X
Uwe Schindler@VEGA:
HTTP/2 200
server: nginx/1.21.0
                            "@context": "http://schema.org/",
date: Tue, 01 Jun 201
                            "@id": "https://doi.org/10.1594/PANGAEA.914629",
content-type: applica
                            "@type": "Dataset",
set-cookie: pansessi
                            "identifier": "https://doi.org/10.1594/PANGAEA.914629",
vary: Cookie, X-PanLo
cache-control: public
                            "url": "https://doi.pangaea.de/10.1594/PANGAEA.914629",
x-cid: 7fc48273de676
                           "creator": [
vary: Accept
                             ₩ {
link: <https://doi.or
                                                                                                                        describes", <https://
9?format=zip>;rel="i1
                                                                                                                        id.org/0000-0002-6952
                                   "@id": "https://orcid.org/0000-0001-8646-2334",
org/0000-0001-6053-2
                                    "@type": "Person",
x-powered-by: Jetty(
                                   "name": "Walter Geibert",
strict-transport-sec
                                   "familyName": "Geibert",
x-ua-compatible: IE=
x-content-type-option
                                   "givenName": "Walter".
x-frame-options: SAM
                                   "identifier": "https://orcid.org/0000-0001-8646-2334",
                                   "url": "http://www.awi.de/ueber-uns/organisation/mitarbeiter/walter-geibert.html",
{"@context":"http://s
                                                                                                                        bs://doi.org/10.1594/
                                                                                                                         "Walter Geibert","fa
angaea.de/10.1594/PAI
                                   "email": "walter.geibert@awi.de"
ter","identifier":"hi
                                                                                                                        er/walter-geibert.htm
                                },
{"@id":"https://orci
                                                                                                                         , "givenName": "Jens".
                             ₩ {
-0002-6952-2494","ur
                                                                                                                         iessen@awi.de"}.{"@id
2330","@type":"Persor
                                                                                                                        d.org/0000-0001-6053-
                                   "@id": "https://orcid.org/0000-0002-6952-2494",
ta E Wollenburg","far
                                                                                                                        staff/jutta-wollenbur
                                   "@type": "Person",
i.de"},{"@id":"https
                                                                                                                         "givenName": "Ruedige
                                   "name": "Jens Matthiessen",
000-0002-4453-9564".
                                                                                                                        stein@awi.de"}]."name
                                   "familyName": "Matthiessen",
al data together with
                                                                                                                        EA","disambiguatingDe
& Environmental Scie
                                                                                                                        ANGAEA","disambiguati
                                   "givenName": "Jens",
arth & Environmental
                                                                                                                        aset comprises previo
                                   "identifier": "https://orcid.org/0000-0002-6952-2494",
omposition of sedimer
                                                                                                                        /396-3: For PS51/038-
                                   "url": "http://www.awi.de/en/about-us/organisation/staff/jens-matthiessen.html",
                                                                                                                         anktonic foraminifer
 230Th excess, 238U,
composition, allochto
                                   "email": "jens.matthiessen@awi.de"
                                                                                                                         uranium isotopes wer
(Element2) via isotor
                                                                                                                        rying and sieving of
le isotopes on a Fini
                                                                                                                        ate preparation devic
te. 230Thex was calcu
                                                                                                                        previously published
                                   "@id": "https://orcid.org/0000-0001-6053-2330",
                                    "@tvpe": "Person".
                                                                                                                             (U)) Universität Bremen
marum
                                   "name": "Ingrid Stimac",
```









Authors

Authors in PANGAEA with ORCID iD:

19.9%







Authors

Datasets / DOIs with at least one author having ORCID:

37.7%

...for data published in 2020 / 2021:

82.3%









Registered Users

Registered users in PANGAEA linked to ORCID:

39.3%

...for new accounts since October 2016: 62.0%,

since 2020: **65.8%**, since 2021: **66.7%**









Feedback?

Thank You!







