Open Climate Knowledge #OCK

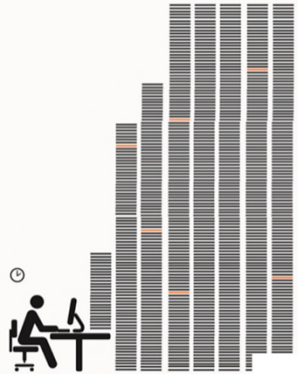
100% open collaborative research for climate change knowledge

Peter Murray-Rust, ContentMine

Simon Worthington, TIB

Join us!

https://github.com/petermr/climate



The days of manually searching through thousands of academic papers are now gone

ContentMine

The Right to Read is the Right to Mine

Anusha Ranganath....
@anusha

Cameron Neylon
@cameronneylon 6:52 PM local time

Charles Oppenheim
@charlesoppenheim

Chris Hartgerink
@chjh 10:52 AM local time

Christoph Steinbeck
@csteinbeck 1:52 AM local time

Christopher Kittel
@chriskittel 10:52 AM local time

Ekta Kumari
@ekta 1:53 AM local time

freyja 10:53 AM local time

Gita Yadav
@gy 1:53 AM local time
I am a Plant Genomicist stuck on Invasive Species

Graham Steel
@mcdawg
Social Media Monitor

Jenny Molloy
@jenny

Joseph McArthur
@joemcarthur

Laura James
@lauraj

Mark MacGillivray
@markmacgillivray

Martina Kunz
@martinakunz 1:54 AM local time

Nikita Wadhwa
@nikitawadhwa2601 1:54 AM local time

Peter Kraker
@pkraker 10:54 AM local time

Peter Murray-Rust
@petermurrayrust

Puneet Kishor
@punkish 1:55 AM local time

Richard Smith-Unna
@rsu
Lead developer

Ross Mounce
@rmounce

Stefan Kasberger
@stefankasberger 10:55 AM local time

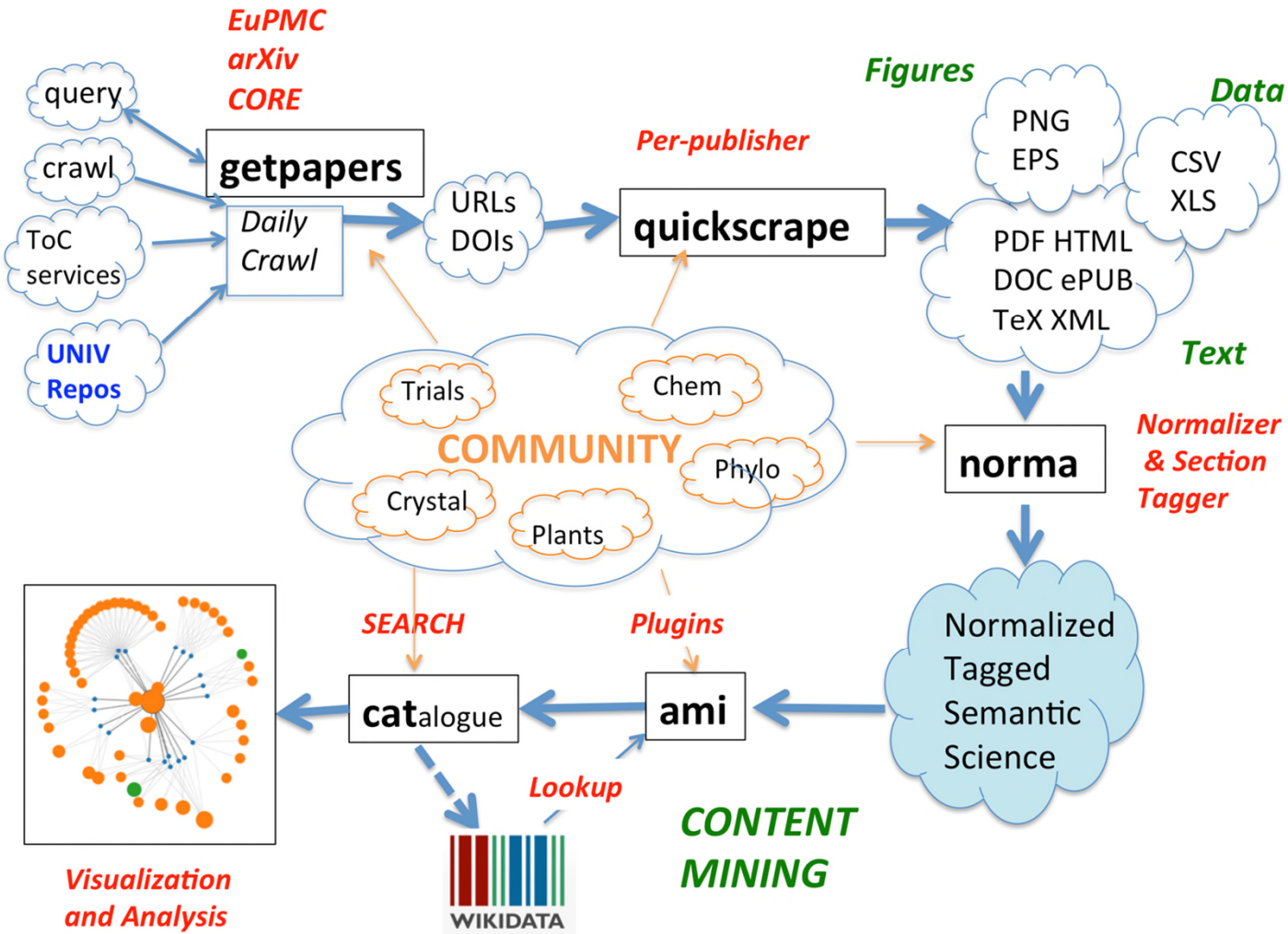
Stephanie Smith-Unna
@steph
Documentation wizard

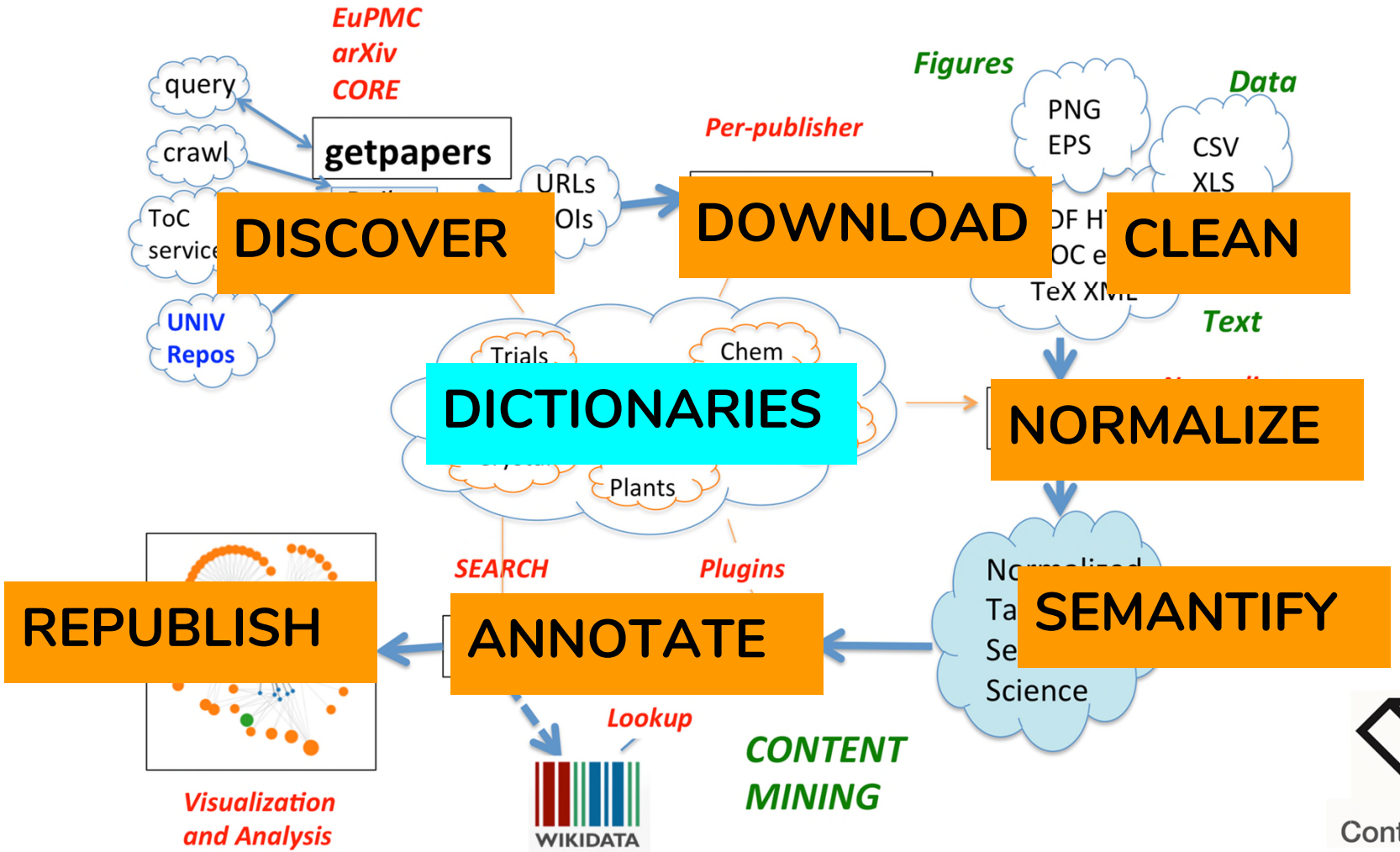
Tom Arrow
@tarrow

ContentMine is OpenLocked Non-Profit http://contentmine.org

Knowledge now comes from many places and involves many people

- Scientific publications
- Preprints, posters
- Theses and dissertations
- Repositories
- Reports, newsletters, press releases
- Domain-specialist sites
- Wikipedia / Wikidata


EuPMC
arXiv
CORE
query
crawl
ToC services
UNIV Repos
getpapers
Daily Crawl
URLs DOIs
Per-publisher
quickscrape
Figures
Data
PNG
EPS
CSV
XLS
PDF HTML
DOC ePUB
TeX XML
Text
Trials
Chem
COMMUNITY
Phylo
Crystal
Plants
norma
Normalizer & Section Tagger
SEARCH
Plugins
catalogue
ami
Normalized Tagged Semantic Science
Lookup
WIKIDATA
CONTENT MINING
Visualization and Analysis
ContentMine


EuPMC
arXiv
CORE
query
crawl
getpapers
ToC
service
UNIV
Repos
URLs
OIs
DISCOVER
Per-publisher
DOWNLOAD
Figures
Data
PNG
EPS
CSV
XLS
CLEAN
Text
Trials
Chem
Plants
DICTIONARIES
NORMALIZE
SEMANTIFY
Science
SEARCH
Plugins
ANNOTATE
REPUBLISH
Lookup
WIKIDATA
CONTENT
MINING
Visualization
and Analysis
ContentMine

How many scholarly articles on Climate change?

```
getpapers -q "climate change" --api crossref -n
info: Found 698927 results
```

CROSSREF

```
getpapers -q "climate change"  -n --api eupmc
info: Found 66838 open access results
```

EuropePMC

http://github.com/ContentMine/getpapers

ScienceDirect has ca 500 K
20% openaccess

Problems

- CLOSED ACCESS.
- DISCOVERY. Few open tools.
- COPYRIGHT.
- Poor/no APIs
- FORMATS. PDF, Word *
- SEPARATION of TEXT, DATA, IMAGES *
- DATA in IMAGES. *
- No semantics, ontologies

* ContentMine software can help.

Knowledge liberation

Currently Open Access rates for research publishing are possibly less than 30%. (Tai and Robinson 2018)

We aim to change that!

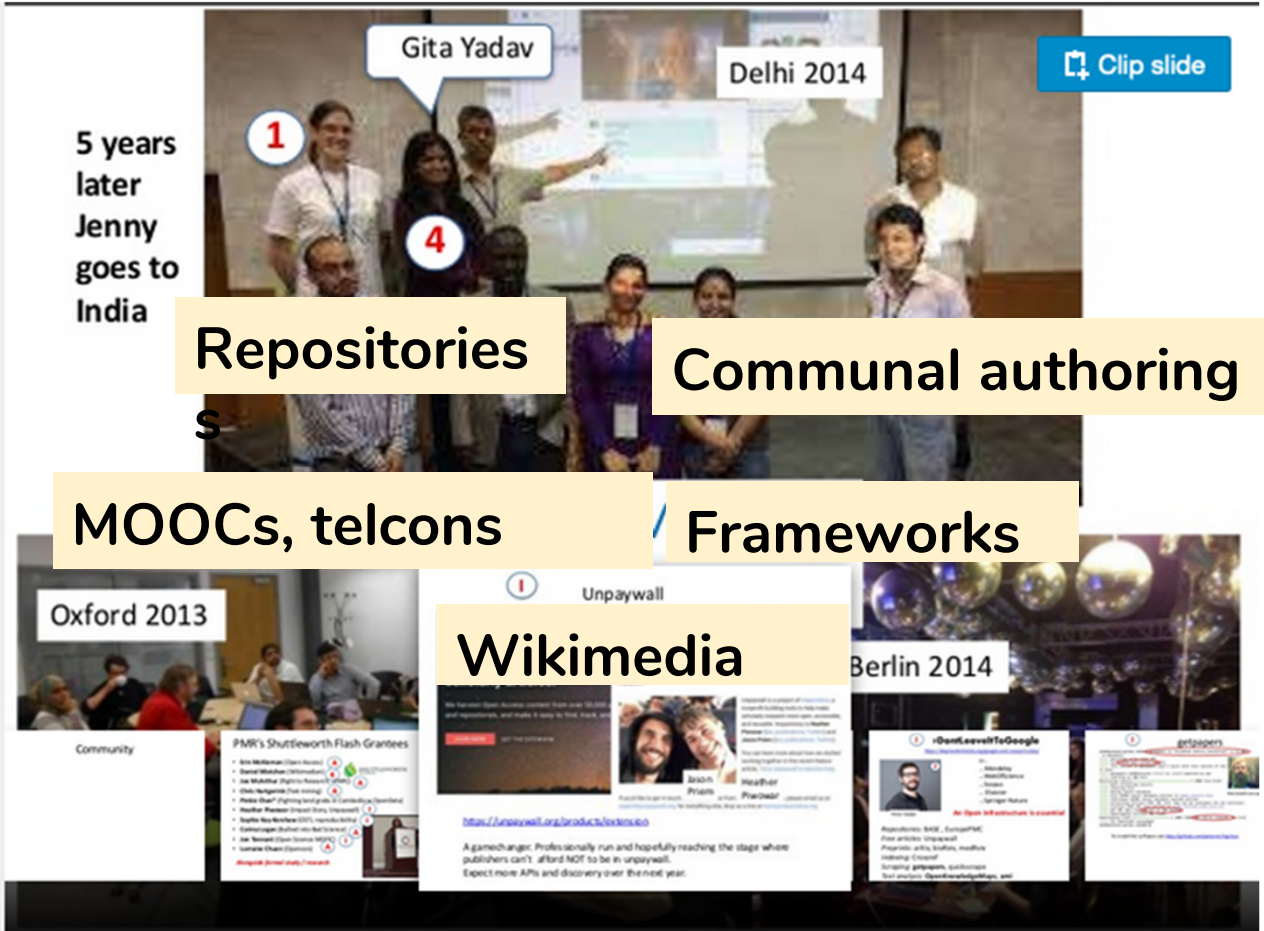
Reference: Tai, Travis C., and James P. W. Robinson. 'Enhancing Climate Change Research With Open Science'. Frontiers in Environmental Science 6 (2018). https://doi.org/10/gf64mq.

EuroCopyright #article 3, 11, 17

Julia Reda, Pirate MEP, running ContentMine software to liberate science 2016-04-16

http://linkme2.net/1vp

Collaborating

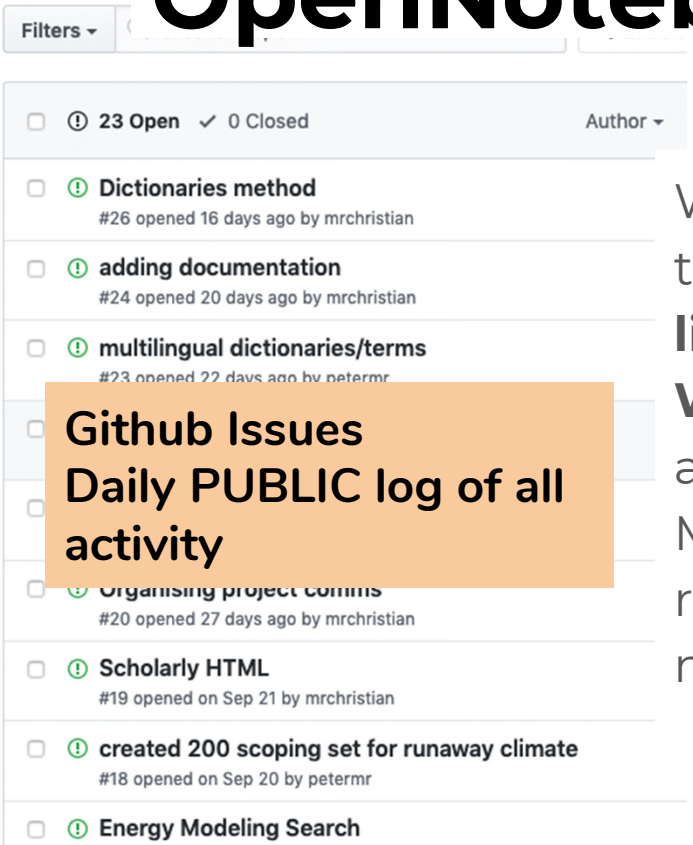


Understanding how to make change

We are doing the following:

- Get up-to-date statistics for open content
- Make a plan for a transition to 100% open
- Create a public repository of all open climate change research publishing

OpenNotebook: Data mining



Github Issues
Daily PUBLIC log of all activity

We have a powerful software OpenNotebook that can search open repositories using **linked data dictionaries based on WikiData**. We can then retrieve metadata and documents and apply analysis with Machine Learning, create collections and repositories, develop Knowledge Graphs, and more.

It's all open and you can use the tools as well.

ContentMining with dictionaries

```
ami-search -p climate --dictionary
 species country funders //builtin
 dictionaries/climate.xml // bespoke


// dictionaries linked to Wikidata
```

OCK – https://github.com/petermr/climate Slides http://linkme2.net/1vp

Scoping review of Climate

Dictionaries ->

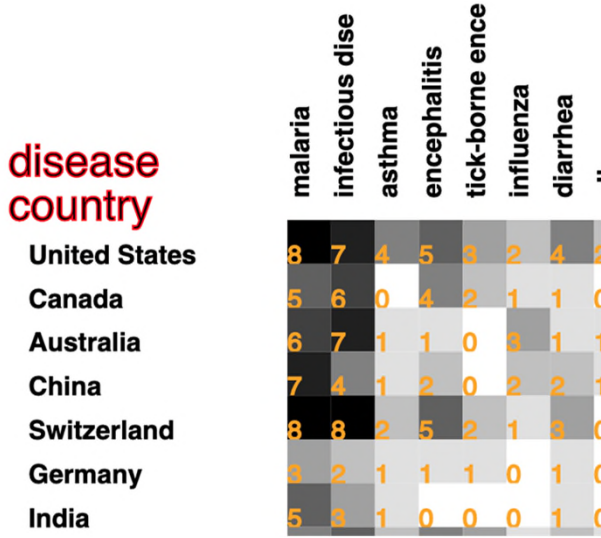
Articles

bibliography	dic:climate	dic:country	dic:disease	dic:funders	species:binomial	word:frequencies
PMC3828158 local	*titl:* The effects of weather and cli ... *date: jour:* *abst:* BACKGROUND: There is much unce ...	climate x 43 weather x 16 Intergovernmental Panel on Climate Change. x 2 IPCC ENSO	Mexico x 15 United Kingdom x 3 Thailand x 2 Puerto Rico. Viet Nam	viral disease tuberculosis malaria	University of East Anglia, x 4 University of California, World Health Organization World Health Organization, Secretaría de Medio Ambiente y Recursos Naturales,	Aedes aegypti x 3	[], x 31 (,–,) x 48 water x 23 change x 27 [], x 36
PMC4105574 local	*titl:* Proximity to coast is linked t ... *date: jour:* *abst:* Psychologists have examined th ...	climate x 101 weather x 3 IPCC x 2 Intergovernmental Panel on Climate Change. x 2 Global warming,	New Zealand x 17 United States x 2 New Zealand, New Zealand. Australia		University of Auckland x 3 Victoria University x 2 Research and Development National Climate Change Adaptation Research Facility, University of Auckland.		change x 77 psychological x 22 belief x 26 distance x 33 New x 32
PMC4113880 local	*titl:* Climate change and human healt ... *date: jour:* *abst:* null	climate x 28 Intergovernmental Panel on Climate Change x 2 weather x 2 IPCC greenhouse gas	Australia x 6 India x 3 Cambodia x 2 Czech Republic United States	infectious disease x 2 cardiovascular disease		Aedes albopictus Ae. albopictus	climate x 28 Public x 24 change x 22 health x 26 Health x 32

OCK – https://github.com/petermr/climate Slides http://linkme2.net/1vp

Scoping review of Climate

articles	bibliography	dic:climate
PMC4675185 local	*titl:* Climate Change and the Evoluti ... *date: jour:* *abst:* Climate change and associated ...	climate x 4 agricultural x 2 Intergovernmental Panel on Climate Change carbon sequestration ecosystem services.
PMC4682990 local	*titl:* Mesocosms Reveal Ecological Su ... *date: jour:* *abst:* Understanding, predicting, and ...	climate x 52 phenology
PMC4744676 local	*titl:* Disciplinary reporting affects ... *date: jour:* *abst:* Climate change is affecting ma ...	climate x 85 Arctic x 11 sea ice x 6 oceans x 5 Intergovernmental Panel on Climate Change x 4



OCK – https://github.com/petermr/climate Slides http://linkme2.net/1vp

Open Science

We're starting with publishing but we aim to cover the whole research cycle, for example:

- methods
- learning
- data
- code

Opportunities

- Training - train-the-trainers,
- Fellowships for Early Career Researchers
- Work with climate centres
- Create dictionaries and ontologies
- Systematic reviews
- Mine Theses and Dissertations

OCK – https://github.com/petermr/climate Slides http://linkme2.net/1vp

Open for who?

Open Science is not only about technical changes, it's also a space for making academia truly universal in practice. OCK is for use by scholars, the public, and for those excluded from knowledge systems — such as the Global South.

- Get involved: what's needed - partners, skills, issues
- Roadmap
- About us: started 9.2019 at eLifeSprint by Peter Murray-Rust & Simon Worthington. The project is an open research project hosted by TIB with *GenR*.

Open for who?

Open Science is not only about technical changes, it's also a space for making academia truly universal in

e by scholars, the public, and for

nowledge systems — such as the

nee

2019

Murray-Rust & Simon W

open research project ho

OCK – https://github.com/pete

Thank you

Open Climate Knowledge

Hashtag #OCK

Join us: https://github.com/petermr/climate