



# Towards Sustainable Scholarly Infrastructures: The Case for CORE

8<sup>th</sup> November 2022

---

Dr. Petr Knoth and Dr. David Pride

Big Scientific Data and Text Analytics Group (BSDTAG)



# Outline

- The case for Open Repositories
- Introducing CORE (CORE.ac.uk)
- CORE services
- CORE Membership



# CORE's role in the OA landscape

Budapest Open Access Initiative (BOAI)  
20th Anniversary in 2022



By “**open access**” to peer-reviewed research literature, we mean its free availability on the public internet, permitting any users to **read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software**, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself.

# CORE's role in the OA repositories network

Confederation of Open Access Repositories



“Each individual repository is of limited value for research: **the real power of Open Access lies in the possibility of connecting and tying together repositories**, which is why we need interoperability. In order to create a seamless layer of content through connected repositories from around the world, Open Access relies on interoperability, the ability for systems to communicate with each other and pass information back and forth in a usable format. **Interoperability allows us to exploit today’s computational power so that we can aggregate, data mine, create new tools and services, and generate new knowledge from repository content.**” [COAR manifesto: The case for interoperability of Open Access repositories, 2011]

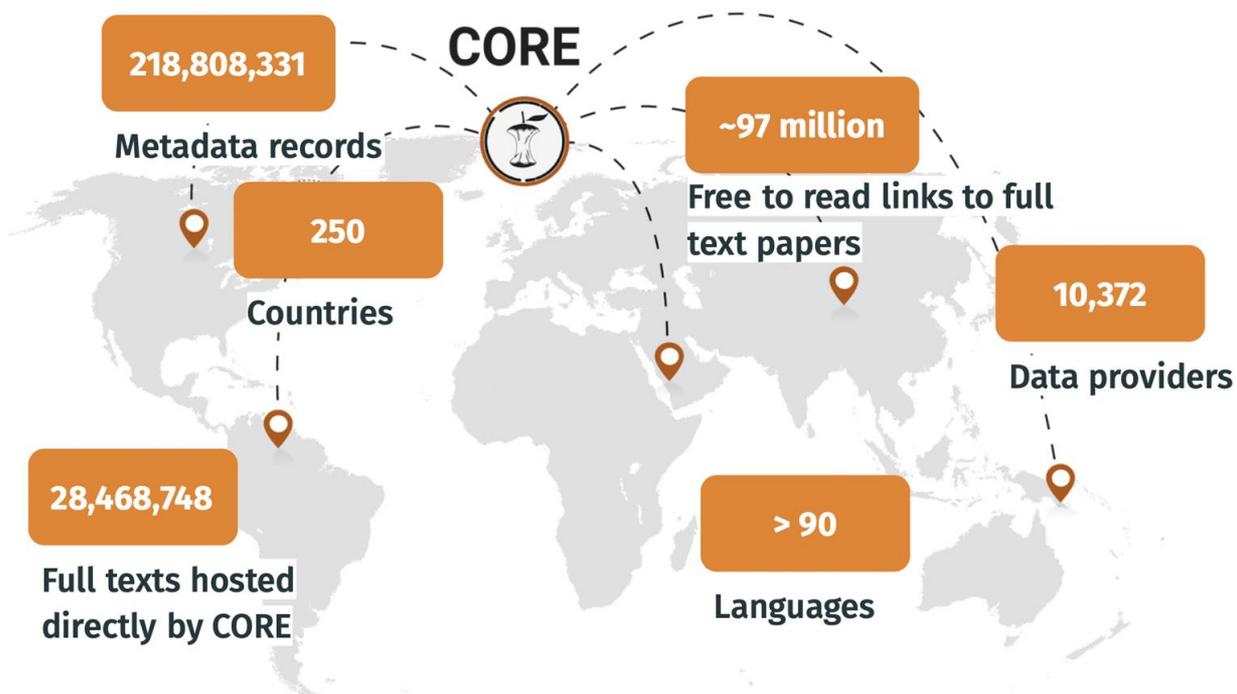
# CORE's mission

CORE's mission is to aggregate all open access research worldwide and deliver unrestricted access for all.

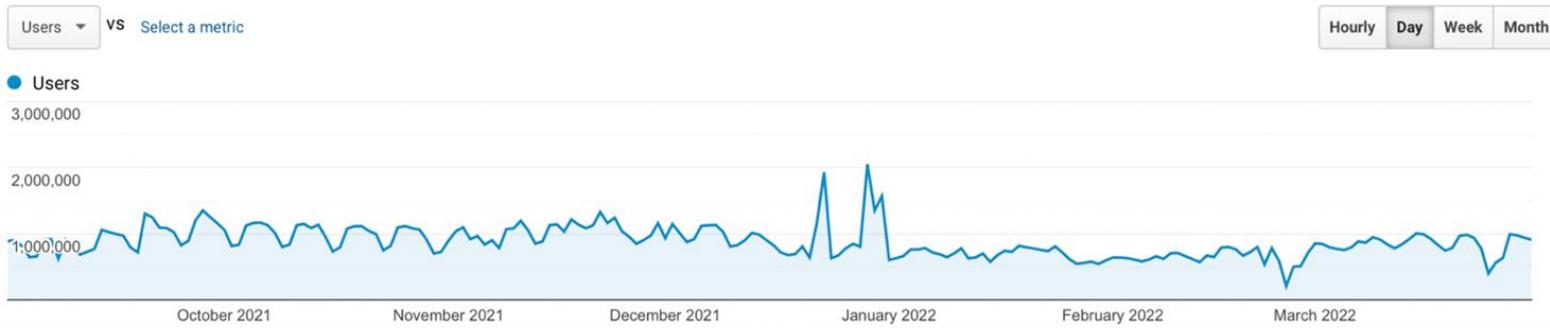
## In doing so, we:

- Support the network of open access repositories and journals with innovative technical, solutions
- Facilitate a scalable, cost-effective route for the delivery of open scholarship
- Enrich scholarly data using state-of-the-art text and data mining technologies to aid discoverability
- Enable others to develop new tools and use cases on top of the CORE platform

# CORE - Global Repository Coverage



# CORE - Usage Statistics 2021/22



Users

186,542,933



Averaging **30 million monthly active users** in 2021/22

# How we rank among services for science and education



## Also visited websites ⓘ

 researchgate.net

 sciencedirect.com

 scholar.google.com

 link.springer.com

 jstor.org

### Other Visited Websites

List of websites frequently visited by core.ac.uk users. Last 3 months analysis

## Topics ⓘ

books education journals

search news internet

software online english

library community science

research reference academic

publications university

34

Science and  
Education

[5th September 2022]

# How does it compare to others?

## Alexa Global Rank [2021]

- CORE (1,373)
- ORCID (5,474)
- Publons (15,459)
- DOAJ (28,387)
- Jisc (133,724) - services sitting under the jisc.ac.uk domain as an aggregate
- SHERPA (154,862)
- BASE (158,709)
- Unpaywall (192,884)
- OpenAIRE (258,719)
- Open Access Button (709,665)



# CORE and The Principles of Open Scholarly Infrastructure

## Governance

- ♥ Coverage across the research enterprise
- ♥ Stakeholder Governed
- ♥ Non-discriminatory membership
- ♥ Transparent operations
- ♥ Cannot lobby
- ♥ Living will

## Insurance

- ♥ Open source
- ♥ Open data
- ♥ Available data
- ♥ Patent non-assertion

## Sustainability

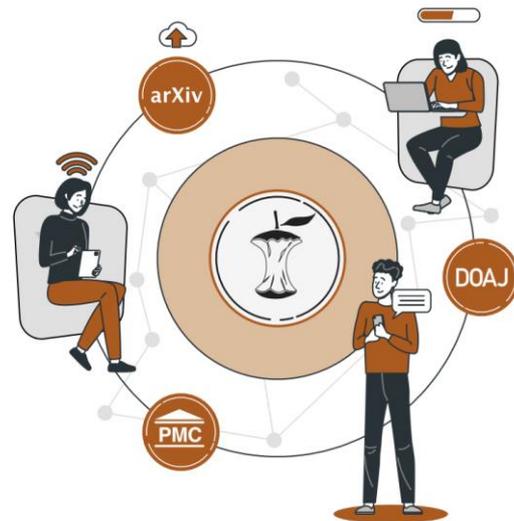
- ♥ Time-limited funds are used only for time-limited activities
- ♥ Goal to generate surplus
- ♥ Goal to create a contingency fund to support operations for 12 months
- ♥ Mission-consistent revenue generation
- ♥ Revenue based on services, not data



**The Principles of  
Open Scholarly  
Infrastructure**

# Become a data provider with CORE

- Increase your contents' discoverability and prevent its misuse
- Make your papers uniquely identifiable and resolvable with PIDs
- Make your content machine readable
- Dashboard for Data Providers / Members



# CORE services

## Content discovery



**Search**



**Recommender**



**Discovery**

## Raw data services



**API**



**Dataset**



**FastSync**

## Managing content



**Repository Dashboard**

**Identifiers**



**OAI Resolver**

# Content Discovery



## Search

The screenshot shows the CORE search results page. At the top, there are filters for 'Find', 'Filter', 'Type', 'Author', 'Language', and 'Publisher'. Below the filters, it displays '14,781 research outputs found'. The main content area shows a list of search results, with the top result being 'Semantics: Towards Full-Text Based Research Evaluation' by D. Hermans, R. Akoth, and A. C. 2016. To the right of the search results is a large image of a bookshelf with the text 'The world's largest open access aggregator' and the CORE logo.

- Full text search for OA content
- Faceted search capability
- What you find is what you get:)



## Recommender

The screenshot shows the Apollo Open Content Mining interface. At the top, it says 'UNIVERSITY OF CAMBRIDGE' and 'Apollo'. Below that, there's a search bar and a navigation menu. The main content area is titled 'Open Content Mining' and shows a list of search results. The top result is 'Semantics: Towards Full-Text Based Research Evaluation' by D. Hermans, R. Akoth, and A. C. 2016. To the right of the search results is a large image of a bookshelf with the text 'The world's largest open access aggregator' and the CORE logo.

- Plug in for repositories
- Relevant recommendations
- Widely adopted: 499 repositories including arXiv.org



## Discovery

The screenshot shows the Digital Object Identifier (DOI) page for 'Translating the self: Ariel Dorfman's bilingual journey'. The page is titled 'Translating the self: Ariel Dorfman's bilingual journey' and shows the DOI '10.1017/S0022268608008083'. Below the DOI, there's a search bar and a navigation menu. The main content area shows the details of the publication, including the title, author, and keywords.

- Plug in for repositories
- Enhanced end-user experience
- Better discoverability for your content

# OAI Resolver

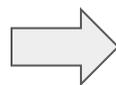


## Resolve an OAI identifier

oai:researchonline.rca.ac.uk:1035

Example: [oai:researchonline.rca.ac.uk:1035](https://oai:researchonline.rca.ac.uk:1035)

RESOLVE



**Royal College of Art**

**Experimental Design: Design Experimentation**

**Tools**

Hall, Ashley, 2011, Journal Article, *Experimental Design: Design Experimentation Design Issues*, 27 (2), pp. 17-26. ISSN 07479360

Information Documents

**Abstract or Description:** This paper was selected for publication in MIT's Design Issues. The research takes an original approach by positioning experimentation as a comprehensive design methodology, rather than using the traditional industrial design approach of employing experimentation as a problem-solving tool within a standard design model. It is an evolution of design thinking on non-linear design methods first developed by Hall and presented to the 'International Association of Societies of Design Research Conference', Seoul, South Korea (2005), and in a paper entitled 'Innovation design engineering: Non-linear progressive education for diverse intakes' presented at the 'International Conference on Engineering and Product Design Education', University of Brighton, UK, which offered a non-linear pedagogy (Hall and Childs 2009) that uniquely supports a diverse interdisciplinary intake. Experimental design is well known in the science domain but very little evidence has been recorded of experimentation in industrial design and its position in relation to work in other science and research domains. Connections are made with theories on research methods, an analysis of case studies and comparisons of literature on experimentation from science disciplines, especially that of Kuhn (1962), Galison (1987), Pasteur's quadrant for scientific research in Stokes (1997) and Borgdorff (2007). Hall makes significant claims in exploring and articulating a model of design experimentation that highlights the differences between scientific and design experimentation. This work was original in describing an experimental design model for the increasing activity in early phases of design development by recording and enhancing knowledge in this important area for future design research and practice. The methods researched in the paper were later used in experimental design workshops in Daegu, South Korea (2011) and Busan, South Korea (2012).

**Official URL:** <http://www.mitpressjournals.org/doi/abs/10.1162/DE...>

**Subjects:** Creative Arts and Design > W200 Design studies  
Creative Arts and Design > W200 Design studies > W240 Industrial/Product Design

**School or Centre:** School of Design

**Identification Number or DOI:** [https://doi.org/10.1162/DESI\\_a\\_00074-Hall](https://doi.org/10.1162/DESI_a_00074-Hall)

**Date Deposited:** 27 May 2012 14:26

**Last Modified:** 09 Nov 2018 14:25

**URI:** <https://researchonline.rca.ac.uk/id/eprint/1035>

# Machine-readable research



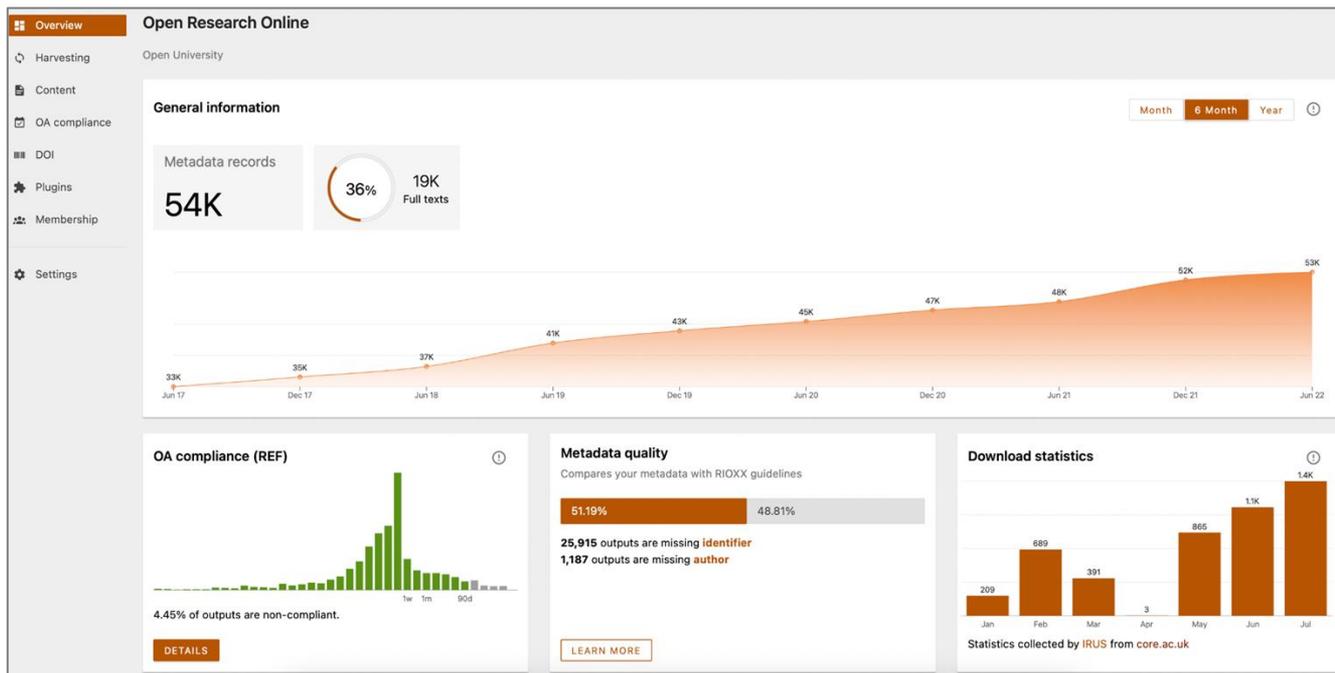
**API**



**Dataset**

- Enables the development of new applications
- Real-time machine access to the world's largest collection of open access papers
- Harmonised access to data from across the network of CORE providers
- Direct machine access to full texts of research papers
- A wide range of application areas: plagiarism detection, fact checking and misinformation detection, research graphs, rapid systematic reviews, detection of research trends, research assessment, identifying suitable peer-reviewers, recommender systems, innovation engineering, compliance monitoring, drug discovery, ...

# CORE Dashboard



# CORE Dashboard

Dashboard Aberdeen University Research Archive theteam@core.ac.uk | [Logout](#)

- Overview
- Harvesting**
- Content
- OA compliance
- DOI
- Plugins
- Membership
- Settings

### General information

Last successful updating: **03/09/2022** Total harvested outputs: **15.03K** 90% 13.52K Full texts

Harvested with **18813** errors

### Harvesting issues

**ALL** ERRORS WARNINGS OTHER

**No download links found**

The metadata provided contained no uris so we are unable to find a full text

**13** records are affected by this issue

**Recommendation**

Ensure dc:identifier contains a url to the full text. Ensure any links on the metadata contain correct uris to the fulltext. Consider linking directly to the fulltext from the metadata.

[DOWNLOAD IN CSV](#) [SEE THE LIST](#)

**Title in the metadata doesn't match the full text**

To ensure we link the correct metadata and full text, we check that the metadata title text is contained within the full text. An exact match is not required, but must be similar. This error is common for scanned items where Optical Character Recognition (OCR) has not been performed. Ensure that the Metadata Title is the same as the document title or including a link to the full text in 'dc:identifier'.

**872** records are affected by this issue

**Recommendation**

We used this as a security measure to not display wrong documents. If the issue is present in many documents, please get in touch with us.

[DOWNLOAD IN CSV](#) [SEE THE LIST](#)

**No valid attachment links found**

We downloaded the metadata, but we were unable to find any full text links.

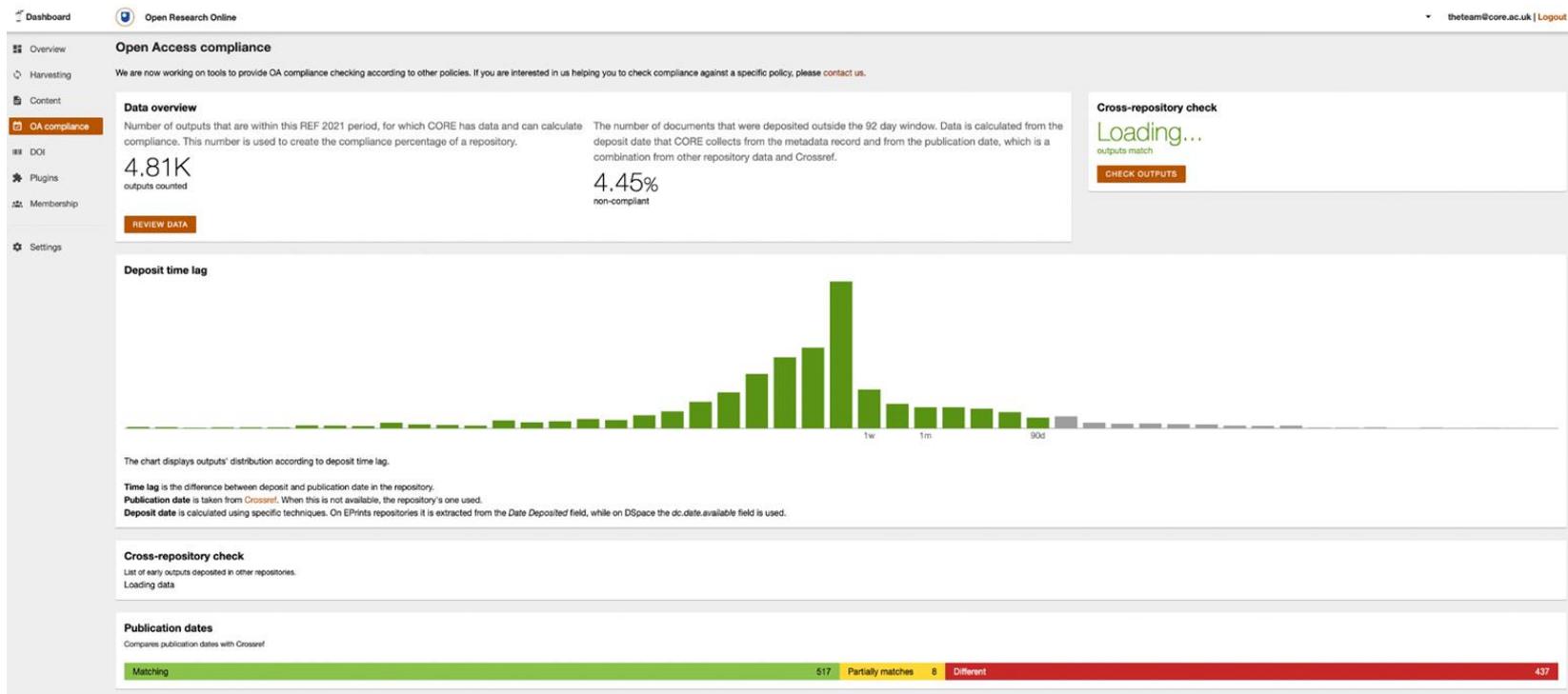
**6** records are affected by this issue

**Recommendation**

Ensure that dc:identifier contains a direct link to the full text. We only support the following file types: .pdf

[DOWNLOAD IN CSV](#) [SEE THE LIST](#)

# CORE Dashboard: OA compliance



# Announcement

- CORE to become an independent Open Access service from August 2023
- CORE will no longer receive direct funding from Jisc from August 2023
- In response, from August 2023, CORE will be operated by The Open University;
  - HEIs from around the world will be able to voluntarily support CORE through a membership programme
  - Members will be invited to on CORE's Board of Supporters.

“We encourage institutions to register their interest in CORE Membership on the CORE website and by doing so, make a strong statement of support for CORE's mission in driving forward open scholarship, as well as benefiting from a range of value-added services.”

**Liz Bal, Director of Open Research Services at Jisc**

# CORE Membership

- A network of data providers who are committed to the ongoing success of the Open Access movement
- Members help shape our development roadmap
- We rely on the generous support of our members to support and sustain CORE



# Three levels of CORE Membership

**Starting**

Free

- CORE Repository Dashboard
- CORE Discovery
- CORE Search
- CORE Recommender
- General support
- Get your OAI identifiers resolved to your repository
- Download statistics
- Repository Discovery Boost:
  - PMC
  - Recommended across the repositories network

**Supporting**

All Starting benefits  
**PLUS**

- Logos for your papers
- Logo banner
- Board of Supporters (1 vote)
- Technical support ( $\leq 5$ )
- API use for your university

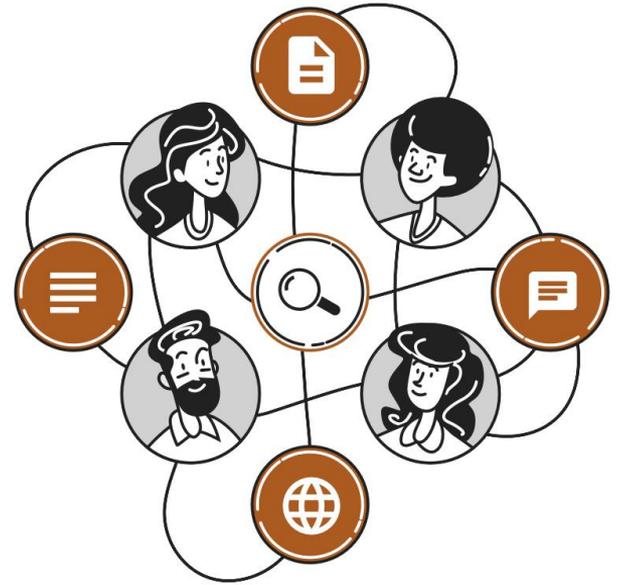
**Sustaining**

All Supporting benefits  
**PLUS**

- Personalised banner
- Board of Supporters (2 votes)
- CORE OA Compliance dashboard
- Unlimited and prioritised support
- Promoted / spotlight / featured articles
- Dataset use for your university
- Hosted or interview style blog post on mission-aligned activities
- Repository health check

# CORE Governance - Board of Supporters

- Part of our commitment to POSI
- Supporting and Sustaining members are invited to join Board of Supporters
- Helping guide and shape the future of CORE



# What others say about CORE ... (1/2)

“CORE has significantly contributed to the global scholarly communications landscape by facilitating a transition from unaffordable and spiralling costs of publisher subscription models to new affordable and equitable systems where no-one, wherever they are, is denied access to knowledge resulting from publicly-funded research”

**Kathleen Shearer, Executive Director of the Confederation of Open Access Repositories (COAR)**

“Over the last decade, CORE has evolved to become a unique global open access aggregator that diligently serves the network of open repositories. CORE promotes and leverages interoperability (OAI-PMH, ResourceSync, ...) across repositories and provides a robust service geared at both humans and machines.

As such, CORE does not only play a central role in providing a seamless layer for accessing OA content from across the distributed repository network, it actually does so in a highly appealing manner. By now, CORE has to be considered as a key established OA infrastructure that is firmly embedded within the OA community and beyond”

**Herbert Van de Sompel, the founder of key Open Standards, such as OAI-PMH and Resourcesync**

## What others say about CORE ... (2/2)

“CORE Recommender is an essential part of the evolving open infrastructure, which surfaces not just the content in our repository, but the entire open literature. Recommender greatly enhances the functionality of our repository”

**Arthur Smith, Deputy Manager of Scholarly Communication (Open Access), Cambridge University**

“CORE is an essential service in identifying open access versions in other repositories and supports monitoring compliance with funder open access policies. It also enables us to get more open access content into our repository.”

**Nicola Dowson, Senior Library Services Manager, The Open University**

“The CORE Repository Dashboard provides an excellent level of feedback and control. It also provides intelligence on errors which may need troubleshooting, thereby ensuring optimum repository health for text mining as well as discovery of repository content.”

**George Macgregor, Institutional Repository Manager, University of Strathclyde**

# What next?

How many data providers have joined CORE?

THIS MONTH

3

LAST YEAR

319

ALL TIME

11K

- Commit to the founding principles of Open Repositories.
- Become a data provider with CORE!
- Register your interest in CORE Membership

<https://core.ac.uk/membership>



**SCIENCE NEEDS YOU**

# Thank you!

<https://core.ac.uk/>

# Panel discussion



George MacGregor

Repository Manager & Technologist, University of Strathclyde. Chair of the Rioxx Governance Group



Torsten Reimer

Dean of the University Library at the University of Chicago



Andy Hepburn

Head of Analysis, Research England, UKRI



Tom Cramer

Director, Digital Library Systems & Services. Chief Technology Strategist at Stanford University Libraries

# Panel discussion

- How can open scholarly infrastructures be sustained?
- Who should contribute and what are some of the mechanisms by which this can happen?