Together Building Open Access, Unlocking Global Knowledge.

October 2025

https://core.ac.uk/governance

NEWSLETTER

COMMUNITY NEWSLETTER

Dear Members,

This year marks a significant milestone for CORE -15 years of COnnecting REpositories, supporting open access, and serving the scholarly community that matters most to us. Looking back, these years reflect sustained commitment, consistent engagement, and a determination to build infrastructure that truly benefits researchers, institutions, and the wider public.

Throughout this journey, we have drawn on history and co-developed solutions alongside you. Since the launch of the CORE Membership programme three years ago, we have seen how invaluable your engagement is in guiding our direction. Your insights, feedback, and partnership have been central to shaping CORE's evolution, and we remain deeply grateful for your ongoing support. It is your consistent involvement that underpins our governance, informs our priorities, and inspires our work.

As we celebrate this 15-year milestone, we are delighted to share with you the developments and enhancements we have built for the community. These initiatives reflect both the progress we have achieved together and our commitment to continue evolving in response to your needs. Looking ahead, our promise is to keep working alongside you, co-developing CORE for another 15 years and beyond, building open access infrastructure that is robust, sustainable, and inclusive.

In this newsletter you will find a roadmap highlighting CORE's development over the past 15 years - a testament to our community's collaborative spirit and innovation that has guided our work.

We look forward to discussing these updates and hearing your perspectives at the upcoming Board of Supporters Meeting on Wednesday, 8 October 2025 at 4:00 PM CET.

Professor Petr Knoth

Head of CORE

With a huge thank you to all CORE members

Looking Back, Stepping Forward: 15 Years of CORE in the Open Access Movement

Fifteen years ago, **CORE (COnnecting REpositories)** began as a PhD project with a simple but ambitious idea: to make open research more accessible, not just for humans, but for machines too. At a time when few could imagine tools like ChatGPT answering questions based on vast collections of literature, it was already clear that the future of scholarly communication would depend on infrastructure capable of delivering information in ways both people and machines could understand. What followed was a decade and a half of learning, building, listening and working with the global repository network community to shape a more open, intelligent, and discoverable future for access to research information.

From its first **Jisc grant in 2011**, supporting a prototype that initially linked just 20 repositories, to today's global infrastructure indexing millions of research outputs, the story of CORE has always been one of scale through collaboration. And of impact born from a shared commitment to openness.

In **2013**, after winning a Jisc commissioned tender, CORE became the **official UK national open access indexing service**. Within a year, its reach expanded well beyond the UK, helping international repositories share content and grow their visibility. It was during these formative years that some of the tools we now take for granted, like the CORE Dashboard and Recommender system, first came to life.

The platform matured quickly. By **2017**, CORE had indexed over **10 million full-texts** and began contributing directly to **policy compliance**, supporting **REF2021** through partnerships with **Research England, UKRI**. These milestones weren't just technical they were signals of trust in CORE's service and stewardship and its value to the **Open Research community**.

CORE's own research, such as the development of CORE-GPT, illustrates how essential its data has become in an era where **AI** is reshaping knowledge discovery. As **machine learning tools** increasingly depend on open, structured, and scalable datasets,CORE's index of over **25 million full texts** provides a critical foundation for experimentation, transparency, and reproducibility in scholarly technology. This highlights CORE's evolving role as both a service and a research enabler that supports the development of thoughtful, responsible tools for the scholarly community.

Our recent initiatives reflect where we're heading next. We've participated and contributed to global discussions on scholarly content such as on r**epository discoverability** and **next generation repositories**, led initiatives to facilitate research reproducibility through archival of **research software** (SoFAIR project), and have always been on the forefront of supporting repositories, such as by being a leading voice on issues surrounding AI bots and repositories. We have also deepened our community model to ensure those who use CORE have a voice in how it evolves through our Board of Supporters.

Looking at **2025**, CORE isn't just marking a milestone, it's stepping into a new chapter with the same curiosity and care that defined its origins. This year, we've launched a dedicated Rights Retentions Statements (RRS), Data Availability Statements and SDG modules, hosted panels addressing the **pressure of AI bot traffic on repositories,** and received multi-year support from Microsoft's Open Data Campaign to continue strengthening global open infrastructure and conduct research into the monitoring of open access content worldwide.

Anniversaries are not only about achievements but also about pausing to thank our community and set our sights on the future. CORE remains committed to open science and we will continue to invest in being the world's most comprehensive index of scholarly literature and a responsible, ethical channel for innovative AI solutions in research and knowledge discovery. This milestone belongs to everyone who has supported, shared data, shaped policy, and guided CORE along the way. **Thank you for walking this journey with us.**

Explore our 15-Year Milestone Roadmap and join us in shaping what comes next.

Foundations of a Movement

Search · Recommender · CORE API · Dashboard v1

Knowledge Media Institute

Petr Knoth started working on CORE as part of his **PhD in KMI**

CORE publishes its vision paper: "*CORE: three access levels to underpin open access*". CORE already has several services:

2010

2011

2012

The Open University

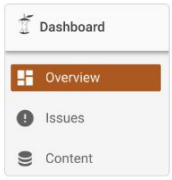
Jisc

CORE (COnnecting REpositories) receives its first grant from Jisc funding 6 person months of work to develop a tool for aggregating metadata from **20 UK repositories** and connecting them via a recommender service.

National Recognition & Service Development

CORE

Repository coverage **expands internationally**; record growth in harvested content.



2013

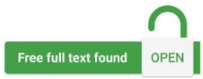
2014

2015

CORE becomes the **official** UK national **OA aggregator** (via Jisc).

Launch of the new CORE Repository Dashboard to improve **repository control for data providers.**

Smarter Infrastructure



Launched a CORE **Dataset service**. Started exploring how CORE could be used for **Open Access policy compliance**.

Launches the CORE **Discovery plugin** and releases **API v2.0** to enable global access for developers.

2016

2017

2018

Dataset

Surpassed **10M** full texts

OR2016

CORE becomes the worlds largest index of open access.

Policy Impact & Assessment Tools

30M monthly active users

25M full texts

Dataset surpasses 25 million full-text records. 30 million monthly active users.

Embracing collaboration with the OA community

CORE launches a **community governance** model.

CORE adopts **POSI** principles.

Membership model

Open Research in the age of AI

- CORE releases its Data Provider's Guidelines and leads a panel on repository discoverability at OR 2024
- CORE leads an international project **SoFAIR** to help address the reproducibility crisis in research by supporting archival of research software and presents this work at **UNESCO**.
- CORE receives substantial multi-year support from **Microsoft's Open Data Campaign**.

Microsoft

2019

2020

2021

CORE selected by Research England /UKRI to check OA compliance in **REF 2021**.

REF2021 Research Excellence Framework

Launch of CORE **API v3**, supporting more scalable and ethical access to scholarly data.

2022

2023

Award-winning paper on CORE-GPT **(Best Paper @ TPDL 2023)** for credible AI-based OA research Q&A. Launching the Deduplication Module. CORE scholarly infrastructure published in **Nature Scientific Data**. CORE launches Deduplication module.

2024

2025

- CORE launches new dashboard modules **(SDG, RRS, DAS)**.
- CORE leads a **panel at OR 2025** addressing the problem of AI bots demand on repositories.

15

Together Building Ope Access, Unlocking Global Knowledge.

15 years of CORE

OPEN REPOSITORIES 2025

Data

The Shape of Comprehensiveness: 13,000 Data Providers and counting.

In the world of science, it's common to talk about big numbers. Citation counts, impact scores, and download figures often dominate the conversation, somewhat controversially. But numbers on their own rarely tell the full story unless they're connected to purpose.

As of **4th August 2025**, CORE (COnnecting REpositories), has surged past the direct data provider mark, a milestone that underscores both our accelerated growth and truly global coverage. It's tangible evidence of CORE's constant growth and truly global coverage. Unlike some similar services, CORE's figure reflects direct data providers, meaning that intermediaries such as DOAJ are counted once rather than tallying every individual journal, offering a clearer and more transparent measure of reach.

At CORE, we actively curate, maintain, and support these data providers, **ensuring they remain operational and fixing issues on a daily basis** not only for our users, but for the **benefit of the global repositories community.** Unlike services that restrict themselves to content with registered (often paid-for) DOIs, we prioritise comprehensiveness and equal visibility of research outputs from all parts of the world. We share this number because it says something vital about the shape of open research and who is included in it. A perfect example is the groundbreaking paper **"Attention Is All You Need"**, which appears in CORE's index via preprint repositories as a seminal which originally didn have a DOI nor ISSN, showcasing that ground-breaking research can be available outside of traditional commercial publishing channels.

Comprehensiveness has always been at the centre of CORE's work. Not because size alone is impressive, but because scale without diversity, depth, and intention is just surface-level visibility. From our earliest days, we've known that if open science is going to serve the world, it has to reflect the world, not just the parts of it that are well-resourced, DOI-registered, or aligned with commercial publishing pipelines.

While many bibliographic databases consider only c**ontent registered by DOI registration** agencies such as Crossref, (one of the largest in the world), **CORE actively indexes a far wider range of research**. As of **4th August 2025**, and in comparison to other private commercial scholarly infrastructure services, including **OpenAlex**, **Web of Science** and **Scopus**, **CORE** works at a global scale, surpassing their reported totals. In fact, based on CORE's internal analysis (2025), around half of the works in our index don't have a DOI at all. This means countless doctoral theses, working papers, local conference outputs, grey literature from underrepresented regions and sometimes seminal papers from key conferences like **NEURIPS**, a prestigious machine learning research conference that doesn't register **DOIs**, find a home in CORE when they might be missing everywhere else.

This is not a minor detail. It's central to how knowledge becomes visible or invisible. Infrastructures that exclude work without DOIs aren't just incomplete; they're making a choice about what kinds of research and researchers get to shape the global conversation.

Data

But CORE also invests heavily in integrating the long tail of research, including smaller, more specialised repositories, each with its own unique attributes, because every piece of research matters. Behind the scenes, this requires a sustained, systematic effort: regularly detecting and identifying new repositories from diverse sources across the internet (beyond reliance on registries like OpenDOAR, which are no longer maintained); uncovering and **diagnosing common issues** that prevent repository content from being discoverable; and communicating with repository managers through a mix of **email notifications** and dedicated **Dashboard tools** for our **supporting repositories**.

Through debugging and testing, CORE actively facilitates the release of vast amounts of scholarly information into the public domain that might otherwise remain hidden. This isn't just a theoretical benefit for example; our work with the UK's USRN has unlocked significant content, and we're now applying similar approaches to **improve repository visibility** in countries like Uganda and Nigeria.

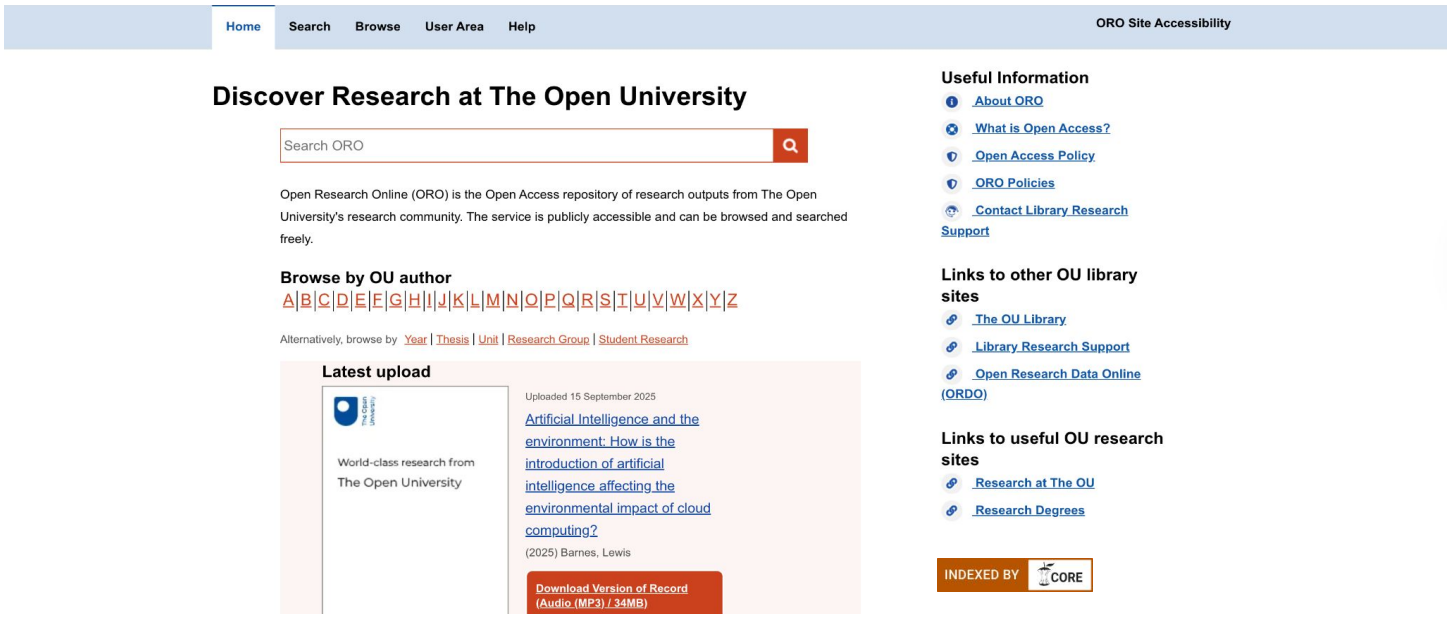
This dual approach, **investing in large-scale harvesting** while directly **supporting smaller and often overlooked repositories** is a key driver of our growth in both data providers and comprehensiveness. Reaching **13,000 data providers** is not simply about scale; it reflects the importance of a model of discoverability that is open, inclusive, and equitable. Unlike commercial services that silo content for their own gain, **CORE works directly with repository communities**, providing services through our **Dashboard** and giving back to ensure **repositories** themselves **benefit** from the process.

We do this not only for CORE, but for the scholarly community as a whole, because when r**epositories become more discoverable, everyone benefits**, including the existing commercial research infrastructures.

As we continue to mark 15 years of CORE, this milestone is part of a larger reflection: progress in open science doesn't come from indexing what's easy. It comes from creating systems that recognise the value of every contribution and from committing to comprehensiveness not as a feature, but as a foundational principle.

To every data provider that's helped shape CORE into what it is today, thank you. This milestone is yours, too.

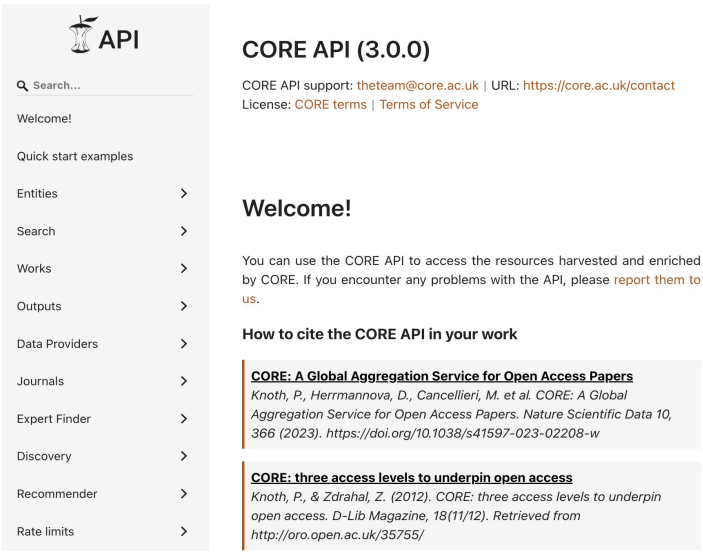
You can use badges on your website to show that your content is indexed by CORE and that you are a part of CORE and Open Research community. More badges



Recent improvements to CORE's DATA services

Improvements to the CORE API documentation - and work on the new

The CORE API documentation has been significantly modified to enhance clarity, user experience, and technical comprehensiveness. Key improvements include **consistent formatting across all entity types** (Works, Outputs, Data Providers, Journals), **real-world usage examples**, and **Python code snippets for each operation**. The documentation also now features a visual representation of the current **CORE data model**. Additionally, the updated API documentation features an **improved query language section** that offers clearer guidance on advanced search constructs, such as logical operators, along with expanded examples for aggregation, free-text search, keyword queries, and phrase matching.



The updated CORE API documentation significantly enhances the developer experience by offering **clear**, **consistent**, and **example-rich guidance across all endpoints**. With clearer API guidance, improved field-level explanations, and actionable examples, researchers, developers, and institutional partners can better understand, navigate, and build upon CORE's infrastructure, accelerating integrations and innovation. The new documentation supports complex use cases while **enhancing API usability, transparency, and user trust** through clearer structure and detailed guidance.

Upgrading CORE's Data Model: Building a Stronger Foundation for Open Research

At CORE, our mission has always been to provide **the most comprehensive and enriched view of research outputs** from repositories and journals worldwide. To continue fulfilling this mission and to keep pace with the evolving needs of the open research community, we are **upgrading our data model** to create a more structured and scalable foundation for the future.

This work is **still in progress**, but it represents a major step forward in how CORE will manage and enrich scholarly data. The upgraded model introduces a **layered architecture** that goes beyond the traditional focus on works, outputs, and data providers, and instead clearly separates data into distinct layers:

- **Raw Files Layer:** Will hold unprocessed data such as XMLs harvested via OAI-PMH, PDFs in the CORE database, and JSON files from other external sources.
- **Initial Database Layer:** Will store all fields extracted from the raw files, capturing **raw, unrefined metadata**.
- **Normalised and Entity Layer:** Will contain **clean, enriched data**, with a focus on **key scholarly entities**, including **scholarly works, outputs, authors, affiliations, data providers, topics, grants, and funders**. These entities will serve as the **building blocks** of CORE's **scholarly knowledge graph**, providing a structured representation of the relationships within the scholarly ecosystem.

This structure will ensure that raw data, metadata, and normalised entities are processed in a **modular and transparent way**, streamlining the flow from harvesting to storage and enrichment. By focusing on these key entities, CORE will be able to offer a much richer and more comprehensive understanding of the research landscape, enabling deeper insights into how scholarship is produced, connected, and funded.

The upgraded model will enable CORE to:

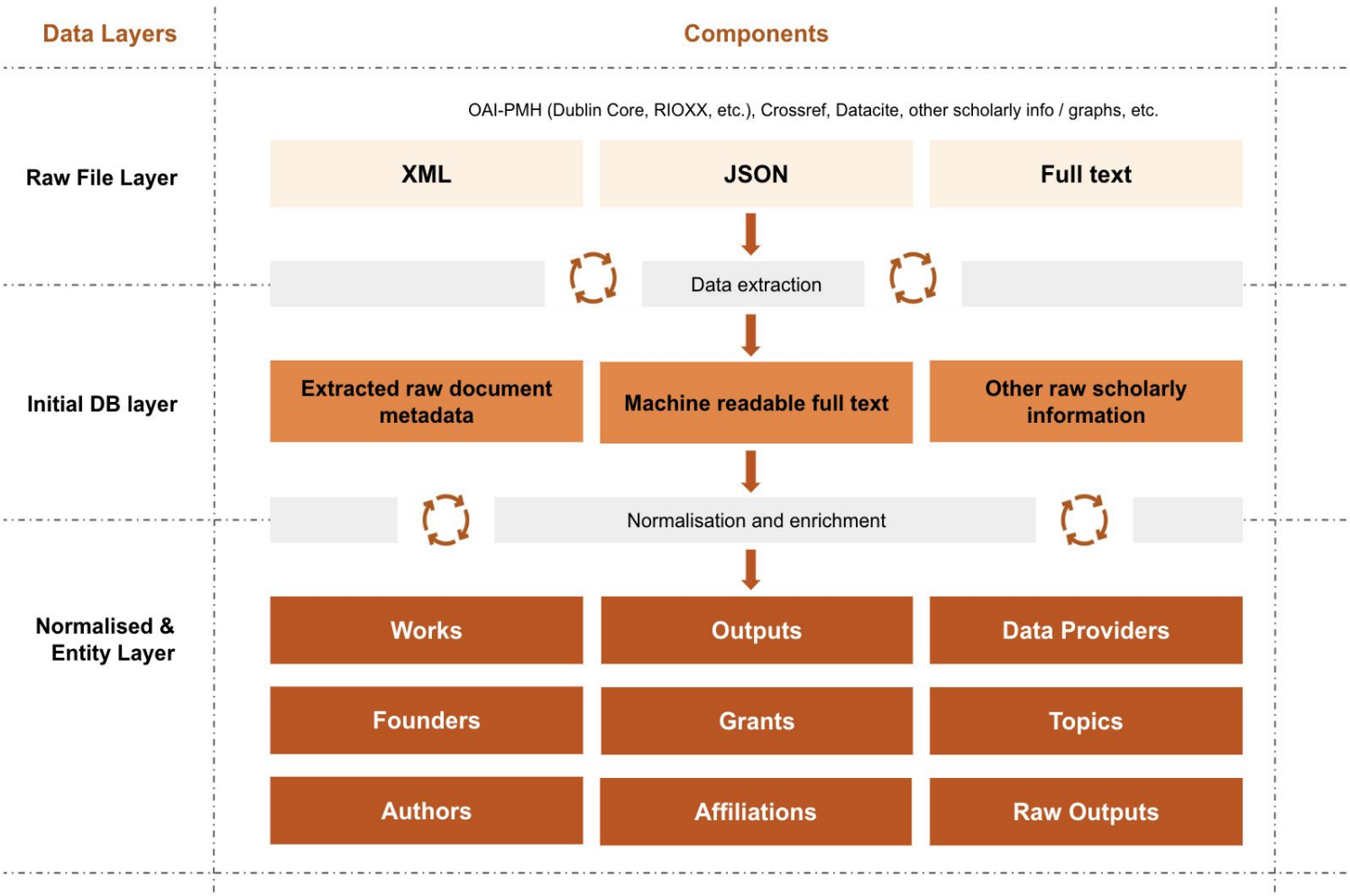
- **Enhance its capabilities** in powering **value-added services for repositories**, such as dashboards, APIs, and recommendation systems.
- **Improve the handling and quality of metadata**, ensuring repositories benefit from better discovery and reporting.
- **Facilitate inferencing and normalisation** over scholarly data, enabling CORE to represent its content as part of an **open scholarly knowledge graph** built from these key entities.
- **Synchronise metadata faster and more efficiently** between CORE and repositories, providing up-to-date and accurate records.

By distributing data across different layers and storage systems, the upgraded model will also provide **greater scalability and adaptability**, ready to support future technologies and services.

Our aim is to ensure that CORE continues to be both:

- A **comprehensive source of research content**, faithfully representing the data harvested from repositories.
- A **provider of scholarly entities that are forming part of a global open scholarly knowledge graph**, driving advanced discovery tools, analytics, and open infrastructure for research.

We anticipate this work to **continue throughout 2026**, as we refine the model and collaborate with our community to ensure it meets the needs of repositories, researchers, and institutions worldwide. Together, we are building a **resilient and future-ready infrastructure** that strengthens the global network of open repositories and empowers the open research ecosystem.



Recent enhancements to CORE Dashboard modules

ORCID

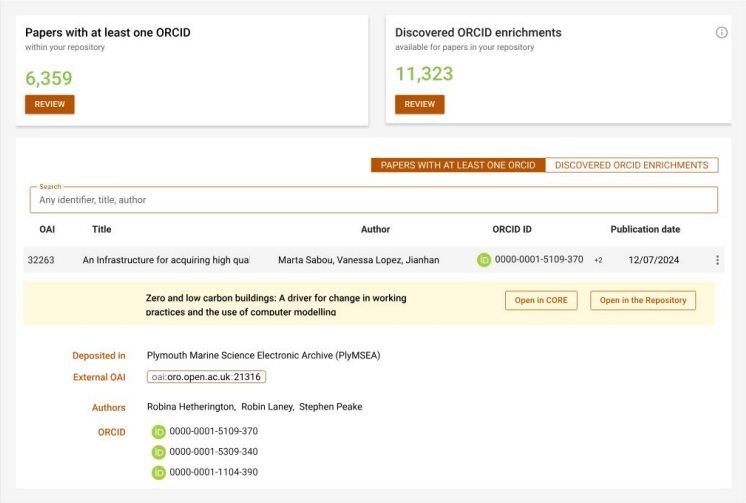
Released

The ORCID tab in the CORE Dashboard helps repository managers to improve author identification and metadata quality. It includes two sections:

Papers with ORCIDs: Shows how many papers in your repository include ORCID identifiers. You can view them per paper or download the full list as a CSV.

Discovered ORCID enrichments: Displays ORCIDs which CORE has found for papers missing them in your metadata, sourced from other providers.

🔗 Read more



Desirable Characteristics Report

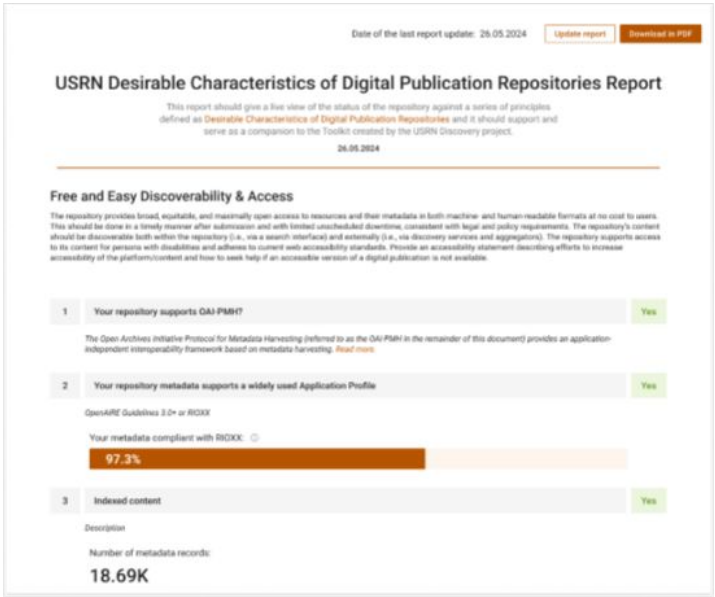
Released

Using the USRN Desirable Characteristics for Digital Publication Repositories, managers can assess how well their repository aligns with best practices.

These characteristics include using unique identifiers **(OAI, DOI, ORCID, ROR**), having a sustainability plan, providing quality metadata, enabling reuse, and ensuring content integrity. The tool offers yes/no or percentage scores, along with actionable recommendations.

It generates a report to highlight areas for improvement. While designed for U.S. repositories, the principles are broadly applicable and valuable worldwide.

🔗 Read more



Introducing the Feedback button on the Dashboard

Released

Feedback

As part of our ongoing efforts to **enhance user experience**, we've added a new feature to the CORE Dashboard: a Feedback button. This addition allows users to communicate directly with our team without leaving the dashboard environment.

The button is designed to be easily accessible, enabling you to **share your thoughts, report issues, ask questions, or comment on any aspect of the dashboard's functionality.** Whether you have a suggestion for improvement or simply want to share your experience, your feedback will go directly to the team responsible for shaping the platform.

Tools and developments

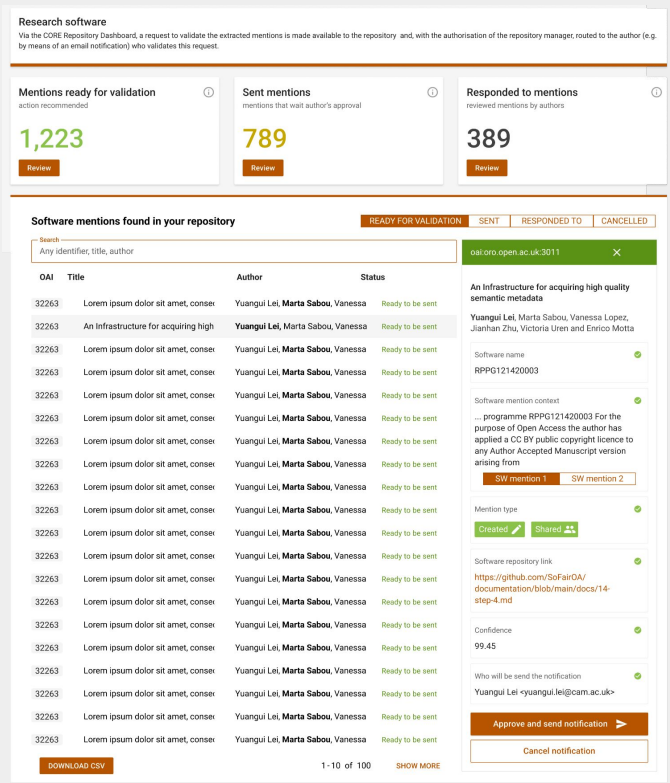
SoFAIR

Project

Facilitating Reproducibility and giving Credit to Research Software Creators

As part of the **SoFAIR** project, funded by the **EU** and **EPSRC**, CORE is piloting a solution aimed at supporting the **reproducibility of research outputs**. The project began in **early 2024** and is now entering its final stages, concluding in **December 2025**. Its goal is to develop a workflow that integrates seamlessly with repositories, making it easier to validate and track information related to research software and other outputs. By embedding this functionality directly into the repository ecosystem, we aim to improve the quality and reliability of data, ultimately supporting open science and reproducible research.

We are pleased to announce the development of a **new module within the CORE Repository Dashboard**, which is currently being piloted. This module will enable repositories to **receive requests to validate extracted software mentions** related to research papers in their repository by the authors of those research papers.



With the repository manager's authorisation, these requests can then be **forwarded to the respective authors.** Authors will typically receive an **email notification**, allowing them to easily review and confirm the accuracy of the information.

To help track progress and engagement, a new **Research Software tab** will display key statistics, including:

- Mentions ready for validation
- Sent mentions
- Responded to mentions

This functionality is **still under active development**, with testing continuing through the end of the project. By December, the pilot will have been fully tested, positioning CORE to explore opportunities for a broader rollout.

CORE attending a key meeting of open infrastructures to address management of research software mentioned in research papers

On the basis of this work, we are thrilled to share that Petr Knoth has been invited to a fully funded workshop organised by Schmidt Sciences in the New York area this October. Schmidt Sciences are bringing together key stakeholders to explore solutions for extracting information about **research software assets** mentioned in scholarly papers. Their goal is to launch a **major funded project starting in 2026** to tackle this challenge.

We hope that future funding from Schmidt Sciences will enable CORE to **scale this work beyond the pilot phase** and **launch it to the global research community**, building on the achievements of the SoFAIR project. CORE's invitation to participate in these discussions reflects the significance of our work and its alignment with global efforts to improve **reproducibility and transparency in research**. We are proud to see CORE's expertise recognised and to play a role in shaping the future of reproducible science.

Recent developments

Integrating CORE functionality with DSpace and DSpace-CRIS

Cooperations

CORE and 4Science share a deep commitment to advancing open access and improving interoperability across the scholarly communication ecosystem. In July, we initiated a strategic collaboration with 4Science, a leading provider of DSpace and DSpace-CRIS solutions, to explore tighter integration of CORE functionality within repository systems.

This work responds directly to feedback from the repository community, which has consistently highlighted the importance of closer alignment between repository platforms and key open infrastructure services. The collaboration is now in the design phase, where we are working together to determine how the integration will operate and which modules it will affect.

While no features have yet been released, the areas under discussion include:

- Metadata enrichment, including persistent identifiers such as DOI, ORCID, and ROR, and Open Access compliance checks.
- Duplicate content detection across repositories.
- Document classification, such as recognising versions, types, domains, affiliations, and SDG (Sustainable Development Goals) labels.

This collaboration represents an important step toward making repositories smarter, more connected, and more impactful. We will keep the community informed as the design develops and look forward to sharing tangible outcomes in the months ahead.

Migration

In progress

CORE is in the process of performing an upgrade of CORE's infrastructure. We are modernising our entire technical architecture, which currently comprises over 20 servers managing all the aspects of CORE.

We are transitioning to a cloud-based infrastructure because we were unable to grow in the physical space and want to provide a more robust, reliable and horizontally scalable solution. We are working on an entirely new architecture that will leverage the full capabilities of cloud computing **whilst working to improve** the reliability and performance our users expect.

Our migration strategy is guided by a few core principles:

- **Scalability and Availability:** We are building a highly available system capable of supporting our continuously growing user base without service disruptions or unsustainable costs.
- **Dynamic Resource Management:** The new infrastructure will be able to support dynamic scaling, allowing us to allocate computing resources dynamically based on current demands for our indexing tasks. Our main goal is always to improve our comprehensiveness of scholarly literature; having a smarter way of allocating resources will allow us to focus where it's needed.
- **Future Innovation Opportunities:** Transitioning to the cloud also allows us to be more **effective** in working with the most recent AI technologies. We anticipate leveraging these capabilities to enhance the quality of scholarly metadata, improve the discoverability of Open Access research, and support the work of data providers in providing the best possible quality research through Green Open Access.

Recent developments

FAIR Certification module

In progress

To enhance discoverability and maximise the impact of repository content, adherence to the FAIR principles is essential. Building on ongoing work within CORE, we are developing a set of measurable goals that indicate how effectively a repository aligns with these principles. Our aim is to establish a FAIR Certification for repositories, offering public recognition of their commitment to FAIR standards and providing external, data-driven validation of their FAIR status.

As we move forward, we're exploring opportunities to offer early access to the FAIR Certification framework. If you're interested in being among the first to engage with this initiative we'd love to hear from you.

If you interested in early trial of FAIR Certification

Register your interest

CASE STUDY

Exciting Update: CORE supports researchers at Erasmus University Rotterdam to develop first LLM for Social Sciences and Humanites

CORE (COnnecting REpositories) is supporting researchers at Erasmus University Rotterdam in the development of **SHARE (Social-Humanities AI for Research and Education), the first Large Language Model (LLM)** designed specifically for the Social Sciences and Humanities (SSH). Unlike existing AI models, which are mostly trained on general web data dominated by STEM disciplines, SHARE draws on CORE's open research corpus to create a model that can engage with the depth and nuance of SSH scholarship.

This milestone is possible because CORE exists as an open scholarly infrastructure, sustained and strengthened by the support of our members. Just as libraries have long enabled researchers to advance knowledge, our community's investment in CORE ensures that academics can now build the next generation of tools powered by open data. In practice, this means that the work you support is directly enabling innovation in fields where it is most needed, from sociology and history to philosophy and economics.

Erasmus University Rotterdam

For the **CORE community** including our **Board of Supporters, Research Network, and Advisory Board** the SHARE project is a vivid demonstration of impact: your contributions are not only keeping repositories visible, but also fuelling breakthroughs that help researchers work faster, smarter, and in new, transformative ways. The project runs until October 2025, and we will ensure its results are widely disseminated to showcase the value of open, member-supported infrastructure.

ANNOUNCEMENTS

Sustaining Open Infrastructure: How Funding Mechanisms Safeguard Open Access

Across the globe, higher education institutions are navigating intense financial pressures. Rising inflation, frozen tuition fees, and the increasing costs of digital services are stretching budgets thin. For repository managers and librarians, these pressures often translate into tough decisions about which services to keep and which to cut. In this context, open access scholarly infrastructure like CORE (COnnecting REpositories) can sometimes appear to be a "nice-to-have" rather than a necessity. But in reality, services like CORE are not luxuries; they are essential infrastructure that keeps research visible, discoverable, and compliant.

Cutting back on open scholarly infrastructure may seem like an easy way to save money, but the consequences are long-term. Universities needs open scholarly infrastructures to provide a viable alternative to prohibitively expensive scholarly databases provided by commercial giants and to maintain digital sovereignty in managing the knowledge assets created using public funds. Governments, the society and the economy need open scholarly infrastructures more than ever as without them (1) our digital sovereignty is compromised (remember what happens when the Internet can be switched off on the Ukrainian frontline or when NASA cannot operate without the support of an unelected oligarch), (2) we cannot deal with threats to our democracy in the (mis)information war (driven by Russia and other state actors targeting liberal democracies)and (3) the long anticipated economic and productivity benefits from AI cannot be realised without these AI having access to trustworthy knowledge embedded in research papers. Our knowledge economy cannot operate well without people having access to knowledge produced by research performing organisations. Open scholarly infrastructures need to deliver on increasing expectations from academics, funders, and policymakers while universities face substantial strain.

The UK as a case study: funding through the UKRI Open Access Block Grant

Thankfully, in the UK, higher education institutions have access to a powerful mechanism that helps sustain open access services: the **UKRI Open Access Block Grant (OABG).** This fund, designed to support the implementation of the UKRI Open Access Policy, can be used flexibly to cover a wide range of eligible costs—including repository support, staffing, and infrastructure.

We know from our own conversations with members that several UK institutions have already chosen to use the OABG to sustain their CORE services. This approach highlights how existing funding streams can be used strategically to protect essential infrastructure. It is a practical, outcomes-oriented way of ensuring that public investment in research also secures public access and long-term discoverability.

Lessons for the global community

The UK's OABG offers a useful case study for institutions worldwide. While funding mechanisms vary by country, the principle remains the same: open access infrastructure is not an add-on but a core enabler of research visibility and impact. In North America, for example, universities have faced six- and seven-figure reductions in collections budgets, underscoring the need to allocate funds strategically to discovery and repository-enabling services. In Europe, Africa, and Asia, libraries face similar pressures, balancing flat budgets against the rising costs of scholarly communication.

Where block grants, consortial funds, or institutional open access budgets exist, dedicating a portion to services like CORE is both compliant and strategically wise. Where such instruments are lacking, the community has an opportunity to advocate collectively for funding mechanisms that treat open access as critical infrastructure rather than optional expenditure.

ANNOUNCEMENTS

Supporting repository teams in difficult times

At CORE, our role goes beyond providing indexing services we work alongside our members to identify practical pathways for sustaining open infrastructure. Whether through existing mechanisms like the OABG in the UK or through collaborative advocacy in other regions, our shared goal is to ensure that repositories remain visible, discoverable, and impactful, even under financial strain.

For decision-makers, the question is not whether institutions can afford to support services like CORE, but whether they can afford not to. Losing visibility layers risks turning open outputs into isolated PDFs. Preserving infrastructure like CORE safeguards compliance, strengthens global discoverability, and supports already stretched repository teams.

As we continue these conversations, including at our **upcoming Board of Supporters meeting,** we invite you to share your perspectives: What mechanisms exist in your region to protect open infrastructure? And where they are absent, how can we work together to build them?

Sustaining open access is a collective responsibility, and together, we can ensure that knowledge remains visible, accessible, and open for all.

REF 2029
Research Excellence Framework

CORE is meeting with UKRI to discuss its role in REF2029

CORE provided data to UKRI for REF 2021 to support Open Access (OA) compliance monitoring. Building on this experience, the CORE team has been working closely with UKRI to explore how this process will evolve for **REF 2029**. These ongoing discussions focus on developing effective mechanisms for sharing and using data to reduce administrative burden and improve the accuracy of compliance reporting.

At CORE, we believe it is essential that institutions have access to the same high-quality OA compliance data we provide to UKRI. By giving institutions clear, timely insights, we aim to help repository teams monitor, track, and demonstrate their compliance with OA policies more easily and confidently.

To make this possible, we are launching a **Working Group of institutions** to collaborate with us. This group will:

- Shape the future development of the **CORE OA Compliance module**, ensuring it fully supports repository teams and their workflows.
- Explore ways to **improve metadata sharing across UK repositories**, so the information needed to track and evidence compliance is accessible, consistent, and actionable.

By joining, your institution will have a direct voice in designing practical solutions that reduce workload, improve data quality, and strengthen the sector's collective ability to meet OA requirements ahead of REF 2029.

If your institution would like to help shape the future of OA compliance support, we warmly invite you to join us in this important collaboration.

Support UK HEI in REF 2029 with CORE – Working Group

Register your interest

Confederation of Open Access Repositories

CORE Joins COAR's New Task Force to Address the Challenges of AI Bots in repositories

The rapid growth of generative AI has brought new opportunities—and new challenges—for the global repositories community. Open access repositories, which provide legally and freely available research outputs, have become a prime target for AI companies seeking training data. Unfortunately, this has led to surging bot traffic, sometimes overwhelming repository systems and threatening the **FAIR principles** (Findable, Accessible, Interoperable, Reusable) that underpin open science.

Earlier this year, CORE surveyed its members about the impact of AI bots. In the March 2025 survey, a significant proportion of respondents reported encountering AI-powered tools accessing their repositories in unexpected or concerning ways, such as excessive scraping, large-scale downloads, or lack of attribution.

This growing concern was also brought to the forefront at two major events: Petr Knoth's presentation at the COAR Annual Meeting and a lively panel CORE co-led at Open Repositories 2025 (OR2025) in Chicago. Both highlighted the urgent need for coordinated, community-driven solutions to balance open access with sustainable repository infrastructure.

Building on this momentum and the strong interest from the community, COAR has launched a **new Task Force on Dealing with Bad Bots**, and CORE is proud to be part of it. The Task Force brings together **key experts from across the world**, representing diverse perspectives, technical knowledge, and regional experiences. It's an incredible opportunity for CORE to collaborate at the forefront of this pressing issue.

The Task Force's work includes:

- Providing **guidance and recommendations** to repositories for addressing both ethical and abusive bot behaviors.
- Exploring **shared services**, such as community-driven whitelists of trustworthy bots or new filtering tools.
- Developing **technical and policy solutions** that support responsible AI use while staying true to open science principles.

You can learn more about the Task Force and view its membership list here.

Having CORE represented through Petr Knoth ensures that the **experiences and needs of our members** are strongly voiced in these global discussions. As a leading indexing solution for open access content, we bring a unique, large-scale perspective on the challenges repositories face and on strategies that work across diverse contexts.

The Task Force has already begun its work and will soon be offering practical advice for repositories, as well as proposals for **new, shared infrastructure** that the community may need to build together. By working collaboratively, we can ensure repositories remain open and FAIR while protecting the infrastructure that underpins open scholarship in the age of AI.

EVENTS

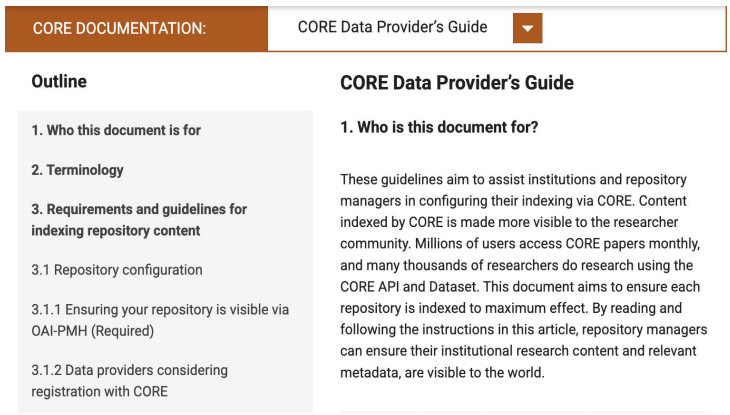
CORE Data Providers' Guide webinar

Link to the guide

Date: Thursday 25 September
Time: from 14:30- 15:45 (UK time)

Join the meeting

Ensuring repository content is **comprehensively indexed and discoverable** in scholarly infrastructures remains a key challenge for the repository community. Even when metadata standards are followed, this alone is often insufficient to guarantee that repository resources can be effectively harvested and surfaced in discovery services. To address this, the **CORE Data Provider's Guide** has been developed as a practical tool to help repositories increase the visibility and accessibility of their research outputs.



The Guide is structured around three essential layers: **repository configuration, metadata configuration, and full-text configuration**. Together, these layers provide a step-by-step roadmap for making repositories "machine-friendly." They cover technical setup, optimisation of metadata exposure, linking and surfacing of full texts, and ongoing validation of repository health. The webinar will also demonstrate practical tools included in the **CORE Dashboard**, which enable repositories to continuously monitor their health, identify issues, and take corrective action to maintain strong indexing performance.

By following these recommendations, repositories can not only improve their indexing within CORE, but also **enhance their interoperability and visibility** across the wider scholarly communications ecosystem.

This webinar will be of particular interest to repository managers, developers, and librarians seeking **actionable strategies** to ensure their repositories are well configured for comprehensive harvesting, discoverability, and long-term health.

CORE at OR2025: Reflections from the Global Stage

READ BLOG POST

CORE brought a series of contributions focused on the responsible use of AI, metadata innovation, and national-level repository coordination to OR2025. Our presentations from the event are linked below:

- SDG Classify: Automating the Classification of Research Outputs into UN SDGs
- Identifying and Extracting Data Access Statements from Full-text Academic Articles
- Managing Machine Access to Open Repositories in the Age of Generative AI
- USRN Discovery Pilot: Increasing the Discoverability of Open Access Content Through a National Network
- How Do You Describe Software in Record Metadata?
- Interoperable Verification and Dissemination of Software Assets in Repositories Using COAR Notify