

# OpenAIREplus Project

Dec 11 - Nov 2014

## Executive Report



Natalia Manola - University of Athens  
Najla Rettberg - University of Göttingen  
Paolo Manghi - CNR-ISTI

On behalf of the OpenAIRE consortium

Feb 12, 2015



## TABLE OF CONTENTS

<b>INTRODUCTION</b>	<b>3</b>
<b>THE OPENAIRE NETWORK</b>	<b>4</b>
<b>POLICY AND INFRASTRUCTURE UPDATE</b>	<b>5</b>
<b>SUPPORT</b>	<b>8</b>
<b>TRAINING</b>	<b>8</b>
<b>DISSEMINATION: WORKSHOPS AND CONFERENCES</b>	<b>8</b>
<b>DISSEMINATION: SOCIAL MEDIA</b>	<b>10</b>
<b>COLLABORATIONS AND LIAISONS</b>	<b>10</b>
<b>LEGAL STUDY</b>	<b>11</b>
<b>TECHNICAL DEVELOPMENTS</b>	<b>12</b>
<b>GUIDELINES</b>	<b>12</b>
<b>AGGREGATION AND VALIDATION SERVICES</b>	<b>13</b>
<b>PORTAL AND END-USER SERVICES</b>	<b>14</b>
<b>INFERENCE SERVICES</b>	<b>14</b>
<b>OA MONITORING SERVICES</b>	<b>15</b>
<b>RESEARCH ANALYTICS</b>	<b>15</b>
<b>ZENODO</b>	<b>16</b>
<b>OPENAIRE APIS</b>	<b>16</b>
<b>D-NET</b>	<b>16</b>
<b>PREPARING FOR THE FUTURE</b>	<b>17</b>
<b>OPENAIRE IN H2020</b>	<b>17</b>
<b>SUSTAINABILITY AND BUSINESS MODEL STUDY</b>	<b>17</b>

## Introduction

The OpenAIREplus project started on Dec 2011 and ended on Dec 2014, building on the previous OpenAIRE project (Dec 2009 - Nov 2012)<sup>1</sup>. OpenAIRE develops and operates a coordination platform for the Open Access deposition and discovery of Europe's research results targeting interoperability on all scholarly communication aspects, from technical to policies to practices. During its 3 years of operation, OpenAIREplus has fulfilled its goals to:

1. strengthen and expand its pan European National Open Access Desks network (NOADs) to all 28 EU member states and 5 associate countries, helping in creating national policies and infrastructures and aligning them with EC's;
2. operate a service-oriented technical platform that interconnects the underlying scholarly communication infrastructure elements, towards the vision of a pan-European Open Research Information System.

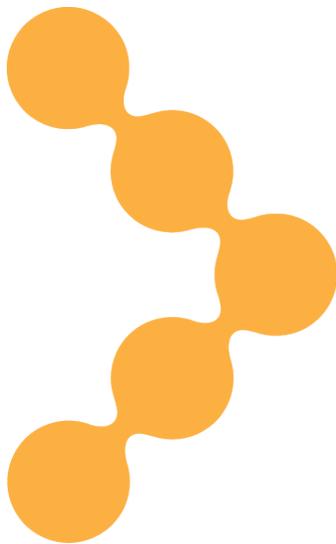
Throughout the EC's FP7 Open Access pilot, OpenAIRE has progressively been building a service-oriented infrastructure, now considered the reference point for Open Access in Europe, establishing a well-recognized brand name within and outside Europe. After five years of continuous operation, OpenAIRE has placed Europe in a leading position for open scholarship, acting as a trusted partner with similar initiatives in the international arena. OpenAIRE, currently entering into the EC's Horizon2020 (H2020), provides **trusted, value added services to a wide range of stakeholders**: researchers to find how to comply with the H2020 and other national or institutional policies, where to deposit their research results and how to interlink them or present them in the right research context; project coordinators and project officers on how to monitor and report a project's scientific outcome progress and its OA policy compliance; institutions and research communities on mechanisms to aggregate their research outcome and impact; data providers and OA publishers to interoperate with emerging e-Infrastructures beyond national boundaries; finally research administrators and the funders themselves to perform research analytics for evidence-based policy making.

Since its first year of operation in December 2010, OpenAIRE has established a production quality interoperable and validated network of more than **580 data providers**, integrating more than **10 mi OA publications**, related to about **25K organizations** and **45K projects** from **3 funders**. It has identified approximately **115K FP7 publications**, produces detailed statistics at project level, and aggregated statistics for research evaluation on Europe's thematic research or geographic regions. Additionally, it harvests metadata on datasets that are related to OA publications or to the supported funders (**6K** so far and counting).

<sup>1</sup> European Commission grants 246686 and 283595

## The OpenAIRE Network

A key element in OpenAIRE since its inception is its **participatory design**, which effectively resulted in the creation of a comprehensive pan-European network, the National Open Access Desks (NOADs), present in 33 European countries. Realizing from early on that no one solution fits all, the NOADs have undertaken the task of approaching stakeholders at the national and institutional levels, trying to understand the specific requirements and idiosyncrasies and aptly influence the local developments on policy, organization and technology. This is a demanding task since it essentially bridges the grand challenges of research infrastructures taking place on the disciplinary, European, global scale to the local practices. These efforts require time, persistence and perseverance and unquestionably need additional strong support by local authorities. Depending on the national status, NOADs have embraced the OpenAIRE activities at different levels of intensity, and it was apparent from day one that many countries with less mature policies and infrastructure sought a community of practice such as the one offered by OpenAIRE. In many cases NOADs played a catalyst role in shaping new OA national policies and aligning them with the EC's, accompanied by implementation aligned with OpenAIRE's.



The **EC's consistent and continuous aid and support** along each step proved invaluable. Even though the FP7 OA pilot was a soft mandate, not strong enough to influence or change researcher practices and mentalities in the broader sense, the infrastructural and distributed solution had already proved that it could serve basic needs of stakeholders in the scholarly communication domain. The July 2012 EC recommendations<sup>2</sup> to member states to align their policies with EC's was a pivotal point that put things in motion: with OpenAIRE services in place, the NOADs closely linked to the newly appointed National Reference Points, and a sufficient number of ministers recognizing the OpenAIRE brand name, a policy shift towards OA in Europe started. This was reinforced by the decisive H2020 mandate and the OpenAIRE placement as a key, effective

monitoring mechanism.

The constantly emerging developments in the scholarly communication environment post challenges that call for continuous adaptations in practices and processes. After five years of operation OpenAIRE recognizes that there are some challenges ahead:

- The NOAD management structure is based on geographic regions, which has certainly provided advantages that mainly stem from similar cultural approaches. As new activities come along, e.g., research data management and its integration with literature, non-

---

<sup>2</sup> [EC Recommendation on access to and preservation of scientific information](#)

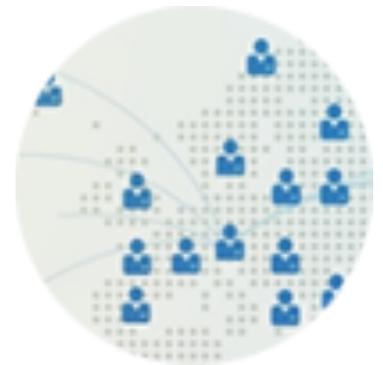
traditional methods of publishing, assessing of scientific impact, OpenAIRE needs to be on the lookout for and implement novel ways for cross fertilization among its members.

- Experience has shown that there is a need for continuous assessment of the network operation and the outreach activities effectiveness. Lean and more efficient communication mechanisms *at all levels* are needed to respond *fast* to new challenges.
- As OpenAIRE reaches its broad vision and more services become available, it is harder to explain its mission and operation to its multiple stakeholders. Clear messages need to be constructed for each stakeholder group and passed on in a direct manner.

### Policy and infrastructure update

---

NOADs are the connector between policy levels and daily OA practice. They are suitably networked, for example many NOADs are in contact with the National Points of Reference<sup>3</sup> for OA in every country, and OA policy groups and sit on relevant national working groups to move the OA agenda forward and working behind the scenes to help shape policy and implementation matters in accordance to the EC recommendations. As a result, NOADs often find themselves in an expert position to disseminate valuable advice to decision makers who are not necessarily well-versed in OA matters.



The following table summarized recent developments in the European OA landscape, many of them influenced by the OpenAIRE work and the NOADs approach.

---

<sup>3</sup> Ministry contacts to align national policies with the [EC Recommendation on access to and preservation of scientific information](#)

Country	Advancements in Policy	Infrastructure Developments
Austria	Major research funder, FWF now has an OA policy for all funded projects	OpenAIRE compliant e-Infrastructures Austria sponsored by Federal Ministry of Science, Research and Economy
Belgium	Installation of an OA working group. The representative from the federal government is also the point of reference to the EC for OA	
Bulgaria	Commission for implementation of an OA plan	National OpenAIRE compliant repository of FP7 publications
Croatia	The new Law of Science and Higher education has a mandate for OA ETDs	A national OpenAIRE compliant repository infrastructure is being set up
Cyprus		All public universities developed OA repositories OpenAIRE compliant
Czech Republic	R&D council released Recommendations on OA, in-line with the EC	OpenAIRE compliant repositories
Estonia	The Research and Innovation Policy Monitoring Program of the Estonian Ministry of Education & Research - analysing current OA trends and providing suggestions on national OA policy, in-line with the EC	
Finland	NOAD liaison with Open Science and Research Initiative (ATT), 2014-2017 by the Finnish Ministry of Education and Culture to follow OA demands of the EC	
Germany	Research Administrators nation-wide very engaged with take up of EC OA mandate	
Greece	Increased awareness of OA among major stakeholders and policy makers	Its emerging national infrastructure is being OpenAIRE compliant
Hungary	HUNOR, association of Hungarian OA repositories promotes repository development	
Italy	Ministry of Education: OA mandate for publications and data, aligned with EC	IRIS information system, similar to CRIS

Country	Advancements in Policy	Infrastructure Developments
Latvia	OA policy drafted by major university, presented to the ministry	90% of FP7 publications integrated into repository
Lithuania	Support from research councils for OA: setting policy principles	National OpenAIRE compliant repository, eLABa
Luxemburg		First repository established, OpenAIRE compliant
Malta		University established first institutional repository
Netherlands	Ministry of Education, Culture and Science committed to OA as of 2014	
Poland	Cooperation with Conference of Rectors of Academic Schools and Ministry of Science to initiate work on policy changes	Development of free repository package: Polish version of DSpace; 383% increase in number of OA repositories
Portugal	Major research funder releases OA policy, in-line with the EC	RCAAP - OpenAIRE compliant infrastructure in place
Romania	National Strategy for Innovation sets OA for all results. Part of Open Government Partnership	
Slovakia	Government plan for OA mandate. Part of Open Government Partnership	
Slovenia	Ministry for Science, Education and Sport set up a Working Group to develop an OA policy, in-line with the EC	Establishment of national OA infrastructure and a national open science portal
Spain	R&D funded projects mirror EC OA mandate	High rate of repositories are compliant
Sweden	The Swedish Research Council has been tasked by the Government to produce national guidelines for open access to scientific information.	
Switzerland	Swiss Science Foundation extends OA policy to monographs	
Turkey	Council of Higher Education has started an OA project	Established many new OpenAIRE compliant repositories

## Support

---

OpenAIRE creates and maintains a comprehensive set of material (FAQs, guides, toolkits) to address different stakeholders requests. As part of the localized outreach it also maintains a Helpdesk ticketing system able to distribute requests to the appropriate experts within the consortium (e.g., NOADs, technical and guidelines team, policy experts). Even though **policy requests are mostly intercepted at the local level** via the extensive NOAD outreach and involvement, other types of questions are intensifying. These are mostly related to repository compatibility and the portal functionalities and results, indicating a service uptake. The helpdesk has proven to be a very efficient tool that allows the OpenAIRE management to reallocate resources where needed.

## Training

---

A range of well attended webinars were held throughout the project, giving a good visibility and a chance to follow-up on contacts and issues. Webinars and other training, especially in 2014 after the H2020 OA mandate announcement, showed a considerable interest that resulted in increased uptake of the infrastructure services.

Topic	Attendance	Countries
Zenodo	106	27
OpenAIRE portal services	51	18
Interoperability with DSpace	105	27
Interoperability with EPrints	42	14
Interoperability with other platforms	38	17
OpenAIRE for H2020 EC Services	100+	EC
H2020 Open Research Data Pilot	70	
H2020 OA to Publications Mandate	62	

## Dissemination: Workshops and Conferences

---

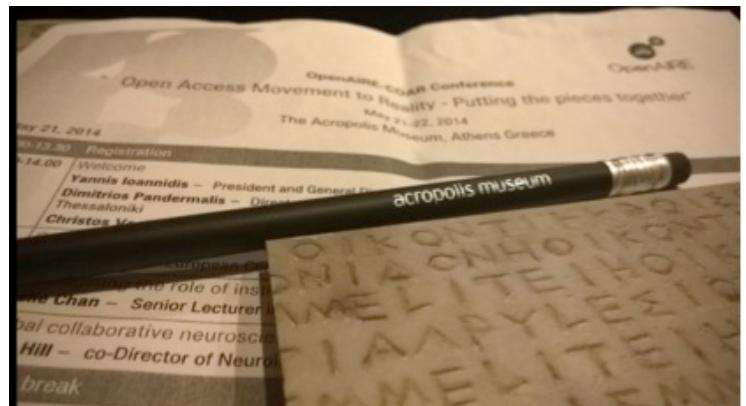
During the three years of the OpenAIREplus project the NOADs have participated in about 1000 national events, which marks a phenomenal outreach activity and illustrates the power of the human network. They have raised visibility of the project's goals and vision via different channels. In addition OpenAIRE organized the following workshops, which allowed the consortium to exchange ideas with external experts on a variety of related aspects:

- Data Linking Policy Workshop, Royal Library in Copenhagen, June 11 2012 with 80 participants. The workshop addressed Research Data policy in the context of linking

publications to research data recognizing that policies and guidelines have to be in place to support research organisations to manage their research data.

- OpenAIRE Interoperability Workshop, University of Minho, on Feb 7-8 with 120 participants, as part of an umbrella Open Access Seminar. It concentrated on interoperability between research infrastructures, namely between literature and data repositories, to fully explore how repositories can communicate with each other.
- Dealing with Data. What's the Role for the Library?, Ghent, on May 28, 2013 with 150 participants, co-organized with LIBER. The workshop aim was to present an overview of possible roles librarians can play in assisting researchers to handle their research data.
- Legal and Sustainability Issues for OA infrastructures, Vilnius, on Nov 5, 2013 with 100 participants. The workshop focused on the two studies currently underway in OpenAIRE which examine two important aspects in running an open access infrastructure: 'Legal and Licensing Issues' and 'Sustainability and Business models'.
- Research Data Management Workshop, a satellite OpenAIRE – COAR event in Athens. Over 60 attendees spent the day discussing aspects of the EC's Open Research Data Pilot, formulating feedback on and the basis for an 'OpenAIRE guide to the Open Data Pilot'. Experts from Canada and the DCC contributed and led the discussions. Awareness among NOADs about the pilot was raised, and a deeper understanding of the role the institution shall play in supporting researchers/project coordinators.

The final project conference, Open Access: Movement to Reality, Putting the Pieces Together, was co-organized with [COAR](#) in the Acropolis Museum, Athens on May 21-22, 2014, with over 160 participants, while an additional 90 tuned in to watch the live stream. The conference was attended by a wide range of stakeholders: NOADs, COAR representatives, institution leaders, librarians, RDM specialists, legal experts, text-mining service providers and publishers, and had two main outcomes: a) it showed the various, interconnected aspects of the OA environment, while providing a glimpse into the future, and b) that OpenAIRE is seen as a key stakeholder in the OA landscape, with increasing number and range of infrastructures showing an active interest for collaboration.



## Dissemination: Social Media

---

OpenAIRE has a considerable presence in the social media as indicated below:

	<b>Twitter:</b> 3.200 followers who provide a high visibility and multiplication effect
	<b>Vimeo:</b> 108K loads and 6K plays of hour workshop and webinar videos
	<b>Slideshare:</b> 138 presentations, 28.255 views and 500 downloads
	<b>Facebook:</b> 660 members regularly communicating and exchanging ideas
	<b>LinkedIn:</b> 200 members discussing topics on implementation mechanisms
	<b>Portal:</b> 3.700 registered users who use the OpenAIRE services
	<b>Newsletters:</b> OpenAIRE gathers information on policies and developments around Europe and the world and sends regularly to 4000 users with a solid, increasing reading of 17-20%.

## Collaborations and Liaisons

---

As the scholarly communication domain is changing and new opportunities and players are emerging, OpenAIRE is out to identify and prioritize key stakeholders, dependencies and evaluate existing or potential collaborations. OpenAIRE's service-oriented infrastructure as well as its brand name have made it a trusted partner within Europe and beyond, with two barriers hindering possible collaborations: i) partners are often not clear about the OpenAIRE role or its long term standing, and ii) OpenAIRE is focussed on European research outputs only, while research is a global endeavour. The former will be addressed soon with the creation of a legal entity, while the latter can only be addressed with OpenAIRE's strong collaborations with similar initiatives worldwide.

Similar initiatives to exploit synergies for the alignment of strategic and technical goals. OpenAIRE is active in COAR initiatives that aim for global interoperability of regional repository networks with CLARA, SHARE, CAS and moreover for technical interoperability with CASRAI, Jisc, EuroCRIS and ANDS. Planned technical pilots will illustrate inter-continental links between research products.

SME's to discuss on how OpenAIRE infrastructure may serve as the basis for their content: researcher networks (ResearchGate, Frontiers, MyScienceWorks, PeerLibrary, Mendeley, REISearch, OA Button group), policy analysts (PPMI), research analytics (ÜberResearch) and CRIS providers.

Open Access journal and data publishers who are looking into solutions for Gold OA and on how to link to data or supplementary data (WileyBlackwell, Pensoft, Ubiquity Press, Copernicus publications, GigaScience, PLoS, Elsevier).

Data repositories and registries to explore better integration with the infrastructure ([PANGAEA](#), [Dryad](#), BioFresh, DataShare, [DataCite](#), [re3data.org](#), [EGI](#), [EUDAT](#)).

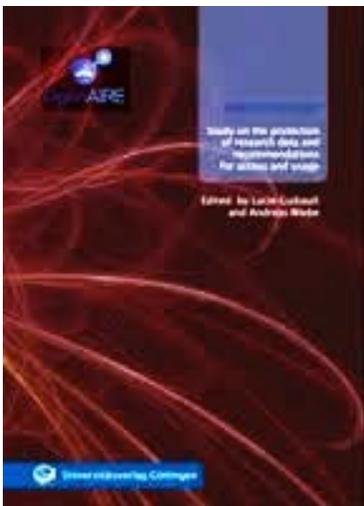
OpenAIRE has a strong presence in various international fora: in RDA, COAR and OpenAIRE have jointly launched an [Interest Group on the 'Long Tail of Research Data'](#) with focus on

institutional research data that falls outside of the 'big data' category; OpenAIRE actively participates in the [RDA/WDS Working Group on Data Publishing Services](#) towards a data broker (resolver from publications to datasets); promotes [Force11 Data Citation](#) related results; liaises with [RMAP](#) and similar projects for literature-data integration.

## Legal Study

---

OpenAIRE analyzed the requirements for the legal protection of research data and types of Open Access licenses available and to which extent some types of usage infringe on the different forms of legal protection, and provided different use-case scenarios. It made recommendations on how the situation as to research data can be improved on a contractual basis between different infrastructures, as well as on the level of legislative initiatives in Europe:



- To respond to the fact that the scientific research exception as presently formulated is inadequate, a new and broader mandatory research exception must be introduced on a European level.
- To achieve legal interoperability of different databases and e-infrastructures, all should license their data under the upcoming CC License version 4.0

The study was published by Göttingen University Press<sup>4</sup>, was widely disseminated and very well received by many stakeholders, indicating that legal issues must be addressed in parallel to any technical developments.

---

<sup>4</sup> Guibault, Lucie, Wiebe, Andreas (Eds) (2013) Safe to be Open: Study on the protection of research data and recommendation for access and usage (CC BY 4.0)

## Technical Developments

The OpenAIRE Information Space is today built by collecting records and content from more than 580 data sources of different kinds, including publication and data repositories, aggregators of both repository classes, and so-called entity registries, e.g., [OpenDOAR](#), [re3data.org](#), and funder data like EC-CORDIS, Wellcome Trust, FCT. The records are harmonized to conform to OpenAIRE vocabularies and formats and used to build a graph of objects of type person, organization, data sources, publication, dataset, and funding. The graph is de-duplicated and enriched via inference services, which work on mining the graph itself and the full-texts of the publications.

With the ever-growing volume of publications, data and their relationships and the need for cleaner and enriched data, OpenAIRE has successfully moved its operations to embrace **big data technologies** (Apache HBASE – Hadoop database). Data intensive mechanisms are now able to run automated workflows to integrate data from diverse sources, to de-duplicate, normalize and integrate aggregated content, to discover features and infer relationships via text mining mechanisms, all processes that ensure a **reliable and trusted data space** ready to be picked up by 3<sup>rd</sup> party providers for academic and research related services.

## Guidelines

---

Guidelines of how repositories expose their content in an interoperable way have proven an essential tool for repository managers and allow such connectivity to take effect while increasing visibility of the repository content. OpenAIRE has thus been a key player in the field over the past 5 years with its “Guidelines for Literature Repositories and OA Journals”. As of Dec 2014, over 580 repositories world-wide have already established a level of compatibility with OpenAIRE’s infrastructure for publications, which as of today covers c. 35-45% of the European repository landscape.

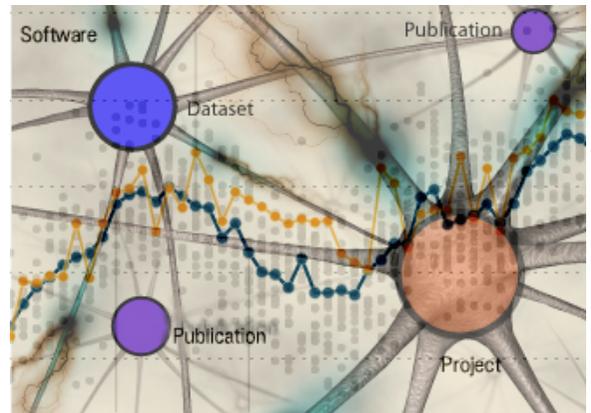
As OpenAIRE is expanding its operations and involves different types of data sources, it has issued three distinct sets of guidelines available at [guidelines.openaire.eu](#):

- ① [OpenAIRE Guidelines for Literature Repositories](#) upgraded to version v3.0 to include funding information from multiple funding agencies.
- ② [OpenAIRE Guidelines for Data Archive Managers](#) adopting the DataCite Schema, extending it to include links to related funding and publications.
- ③ [OpenAIRE application profile for CERIF-XML](#), co-developed with the EuroCRIS group, to allow CRIS systems to export their metadata in an OpenAIRE compatible format.

The guidelines for data repositories uptake is currently rather low, indicating the unclear and uncommitted environment for data standards. OpenAIRE aims to intensify its outreach efforts to create a truly pan-European interoperable network of repositories for all types of scientific results. These processes have already made it clear that efforts must align with co-current international initiatives: [DataCite](#), RDA, ANDS, EuroCRIS, COAR and CASRAI.

## Aggregation and Validation Services

OpenAIRE has expanded its open access infrastructure to include bibliographic metadata to all kinds of open access publications and not only the ones from EC's FP7 programme. With a merge of DRIVER's 6.5 mi publications into OpenAIRE, the current infrastructure harvests 10 mi publication metadata from more than 580 institutional, thematic repositories, national aggregators and OA journals. An increasing portion of them are linked to funding information, showing a clear interest of all stakeholders for a linked research environment.



Regarding datasets, OpenAIRE harvests metadata from data repositories if they are related to an open access publication in OpenAIRE or linked to a project. The current number is around 6K, harvested from a handful of repositories, mainly Zenodo and [PANGAEA](#). The OpenAIRE Guidelines for Data Archive Managers define these relationships, which makes their dissemination, uptake and implementation of high priority.

OpenAIRE also harvests from a variety of so-called entity registries, web-accessible data sources providing authority lists of entities needed to contextualize and enrich the aggregated metadata. Examples are [OpenDOAR](#) and [re3data.org](#), which maintain registries of publication and data repositories worldwide; CORDIS – the EC database that includes all FP7 and H2020 funding information, Wellcome Trust – the UK funder whose data is retrieved from Europe PubMed Central, and FCT – the Portuguese government funding agency (currently in pilot phase).



The validator service, a valuable tool for repository managers, is a rule-based web service, which is used within the OpenAIRE workflows or as a stand-alone service. It can be configured to validate all types of rule sets as these are defined in the OpenAIRE guidelines for literature repositories, data archives or CRIS systems. Continuous developments have been applied to integrate into OpenAIRE workflows, apply integrity checking, increase its performance and make the results presentation more user friendly.

## Portal and end-user services

---

The OpenAIRE portal has undergone through a complete revamp with a new look and feel. Services and workflows for end users have been re-evaluated and modified, due to a series of usability assessment reports. In addition to the information pages, the helpdesk and search/browse functionalities, the portal includes:

- Visual links to [OpenDOAR](#) and [re3data.org](#) that allow researchers to find out where to deposit.
- Landing pages for publications and datasets that display inferred information (ids, subjects, related publications and datasets, references, similar publications), provenance information (links to repositories, journals and aggregators) and usage statistics when available.
- Aggregated results and statistics per project, providing a quick view to project coordinators and project officers.
- Claim services that take end users through a series of guided steps to link objects within or outside the OpenAIRE information space (the latter via the use of 3<sup>rd</sup> party APIs from CrossRef, [DataCite](#), ORCID).

## Inference Services

---

The Information Inference Service (IIS) is a distributed and parallel framework that enables big data workflow-oriented batch processing, effectively allowing OpenAIRE to import content (text, pdf, XML) from a variety of distinct data sources and apply any arbitrary set of parameterized text-mining components. Current OpenAIRE mining components include:

- project extraction for FP7, WT, FET, EGI, FCT and other funding information;
- data citation extraction based initially on links to [DataCite](#) through the identification of DOIs and other similarity algorithms;
- content classification, currently working with training sets from arXiv, WoS, MESH and Dewey classification schemes;
- publication references extraction;
- publications similarity identification.

All mining components are able to work within the IIS framework to batch process millions of incoming documents, but are also published as web services (<http://mining.openaire.eu> with credentials openaire/openaire) to be used by data or service providers to enrich their content.



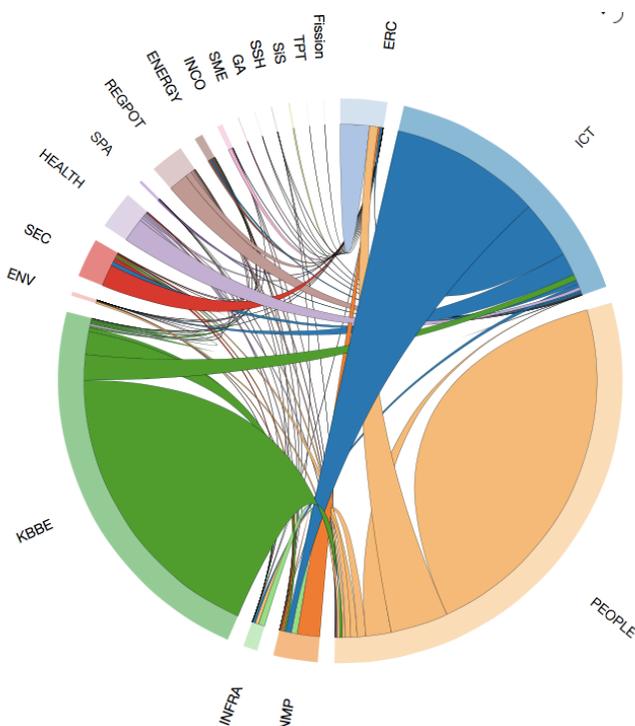
## OA Monitoring Services

OpenAIRE has developed a set of monitoring tools that allow the *on-demand* definition, processing and visualization of its aggregated results. Depending on the enrichment and status of the underlying data, data are aggregated, joined and queried appropriately to produce statistics that evaluate the OA status and progress at the institutional, national, funder or European level. These statistics can then be delivered and visualized to the OpenAIRE or other 3<sup>rd</sup> party portals/services.

## Research Analytics

An intelligent and scalable probabilistic framework has been developed to analyse the evolution of (unobserved) topics of document collections over domain or geographic regions or time. A prototype service analyses massive collections of documents and related metadata and infers interesting or hidden patterns, groups, similarities and latent interrelationships within and across different data types such. Specifically the service

- Discovers **multi-modal "topics"** that tie together research areas with projects, mesh terms, research areas and publication venues characterizing the collected scientific outcome.
- Reveal **"content" based clusters** and examine **similarity patterns** to identify hidden groups, structure and communities.
- Creates **scholarly communications and research map** of research objectives, activities and thematic priorities and assesses different funding programme coverage.



An accompanying *interactive* web visualization tool allows end users (policy makers, research administrators, funders) to view maps and object graphs of past and current research related objects. that essentially help them identify patterns and determine relations among projects and funding streams. The tool offers drill down functionalities allowing them to focus on specific interests. Multiple experiments with OpenAIRE enriched metadata from [arXiv](#) & [Europe PubMed Central](#) datasets have illustrated the values of these services and their potential use by decision/policy makers.

## Zenodo

---

The Orphan Repository for publications was transformed within OpenAIREplus into [Zenodo](#) for **all research artefacts**, including data. The rebranding was well received since it was perceived as more positive and welcoming with a strong identity. More significant than this though was the scope increase to include data, since it was found that whilst researchers had many alternative institute or subject repositories for their publications, there were very few alternatives for data, especially for datasets of any significant size. Therefore Zenodo has received a lot of interest, and coverage in the social media, not only from the smaller institutes looking to move to a hosted model, but also larger institutes seeking trusted long term partners. There is a strong requirement to integrate into other services, ranging from publisher workflows to laboratory toolkits, which Zenodo has been able to meet by exposing a rich API. The most rapid growth area is seen to be in software preservation, thanks to the integration of Zenodo with GitHub, which makes it easy to share and cite software, and it seems that researchers are much more at ease with this than with data, due to the familiarity with the open source methodology.



## OpenAIRE APIs

---

The OpenAIRE APIs were developed, documented and published at [api.openaire.eu](http://api.openaire.eu) allowing developers to access the metadata information space of OpenAIRE programmatically. Work was carried out in collaboration with the EC IT department in order for OpenAIRE to export project results to the EC's participant portal in order to act as a reporting tool for H2020. The API development is work in progress and the OpenAIRE team plans to invite external communities to discuss their requirements for further extensions and uptake.

## D-NET

---

OpenAIRE is a running instance of the open source infrastructure framework [D-NET Software Toolkit](#) which assists developers in the construction and maintenance of aggregative data infrastructures. D-NET is a service-oriented framework where a rich set of services for the collection, processing and provision of metadata and files can be customised and pipe-lined to implement the internal workflows of data management applications. It has currently been installed in National repository aggregators ([Spain](#), [Argentina](#), [Poland](#)) and is adopted in a number of European projects beyond OpenAIRE ([EFG](#) – The European Film Gateway, [ESPAS](#) – Near earth data infrastructure for eScience, [EAGLE](#) - The Europeana Network of Ancient Greek and Latin Epigraphy Data Infrastructure, [HOPE](#) – Social History data infrastructure).

## Preparing for the future

### OpenAIRE in H2020

---

OpenAIRE has produced a [first round of briefing documentation](#) for project coordinators and EC project officers. This has happened in consultation with the ERC. The OpenAIRE team has also started to answer a number of requests via the Helpdesk and its social media channels, about EC funds for Gold OA, and the Data Pilot.

After close collaboration with the EC's backend IT services, OpenAIRE is ready to serve project results reporting in H2020: OpenAIRE identifies research publications and data in its extensive network of data providers, while an EC IT service pulls these reports into its databases. This places OpenAIRE as a key European infrastructure enhancing its visibility and credibility.

The upcoming phase of OpenAIRE (EC funded project [OpenAIRE2020](#)) will involve a variety of activities: address the Open Data Pilot of the H2020; collaboration with national funders to reinforce the infrastructure's research analytic services; an APC Gold OA pilot for FP7 publications; novel methods of review and scientific publishing; a study on scientific indicators related to open access; legal studies to investigate data privacy issues relevant to the Open Data Pilot; international alignment with related networks.

### Sustainability and business model study

---

OpenAIRE commissioned the Athens University of Economics and Business to create a sustainability and business model. The main goals of this study were to i) prioritize the stakeholders with terms of **Willingness To Engage (WTE)** and **Willingness To Pay (WTP)**, ii) provide monetary costs/benefits and revenues for subscription-based sustainability through a comprehensive Cost Benefit Analysis, and iii) give recommendations on how to capitalize on spillover effects and protect the OA nature of OpenAIRE. The results of the study, some of which will form the basis for the development of the OpenAIRE legal entity in 2016, have shown that:

- WTE: University Research Offices seemed to be mostly interested in OpenAIRE's operations, while researchers were less knowledgeable or willing to participate.
- WTP: A subscription model could be a viable option for OpenAIRE, as preliminary results show the WTP for specific services (e.g., interlinking of publications and data) is greater than the corresponding WTE.
- Long term effects: The *simulated* potential R&D effect from the existence of OpenAIRE suggests even larger net social welfare benefits in the long run, with a Benefit/Cost ratio of about 70 for 50 years, going upward for subsequent years.