

NEXUS



# Community Call

APR 20, 2023

**Guest:** Katie Bannon Sr Information Officer



# Agenda

- Intro
- The World Bank MONITOR Dashboard
  - The Monitor pipeline: from PROVIDE to MONITOR
  - Institutional needs, requirements & the MONITOR experience ***Katie Bannon, The World Bank***
- New and upcoming MONITOR Indicators
- Q&A

# INTRODUCTION

---

# WHY MONITOR?



## Know Thyself

Open Science Uptake

Resources

Research Output

Collaborations

Visibility

Impact



## Understand Thyself

Pathways

Opportunities

Insights



## Position Thyself

Decision making

*Reporting*

*Story telling*

# Why OpenAIRE **MONITOR**?

'Not everything that can be counted counts, and not everything that counts can be counted'

*(attributed to Albert Einstein)*

- **Relevance for the community**
  - Co-develop indicators that make sense to all
- **It's all about open science and open data**
  - Inclusiveness, transparency and replicability
- **Coverage**
  - Beyond publications: Research Data, Software, Other Research Products
  - Linked science
- **Fully embedded in EOSC infrastructure**
  - Starting from content providers to included metrics

# How? Methodological Approach

Tracking and evaluating research activities and open science uptake must be

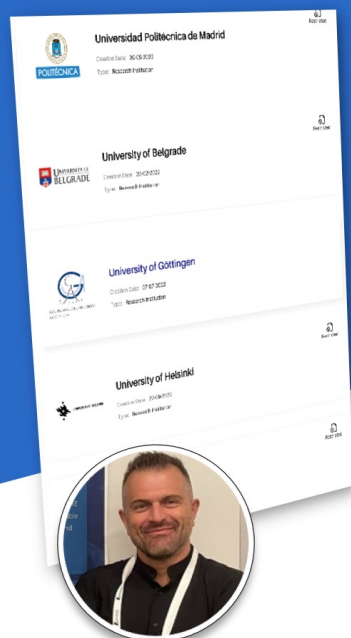
- **data-driven & relevant** – use key aspects of big data available
- **comprehensive & granular** – 360° view across multiple facets of research activities & their links, down to a fine level
- **automated & timely**
- **sustainable** – tracking long-term & repeatedly
- **open, transparent & replicable** → trustworthy



# What? Dashboards on Demand

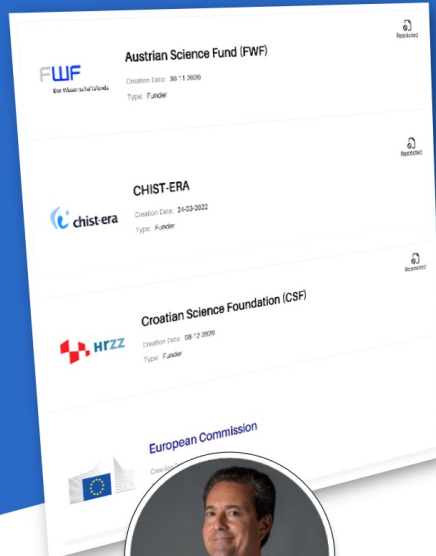


## Institutions



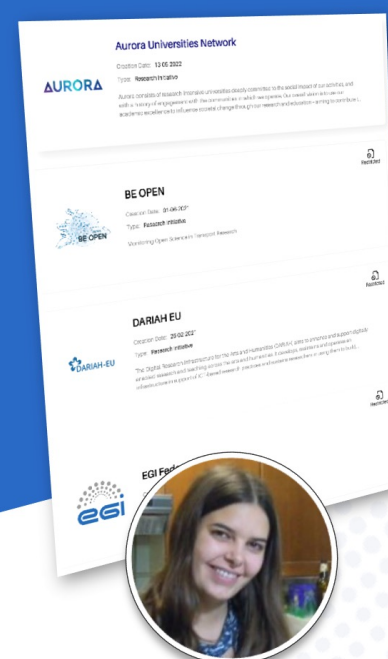
Your OpenAIRE expert  
Leonidas Pispiringas

## Funders



Your OpenAIRE expert  
Harry Dimitropoulos

## Research Initiatives



Your OpenAIRE expert  
Alessia Bardi

## Countries



Your OpenAIRE expert  
Ioanna Grypari



**THE WORLD BANK**

# MONITOR DASHBOARD

From PROVIDE to MONITOR

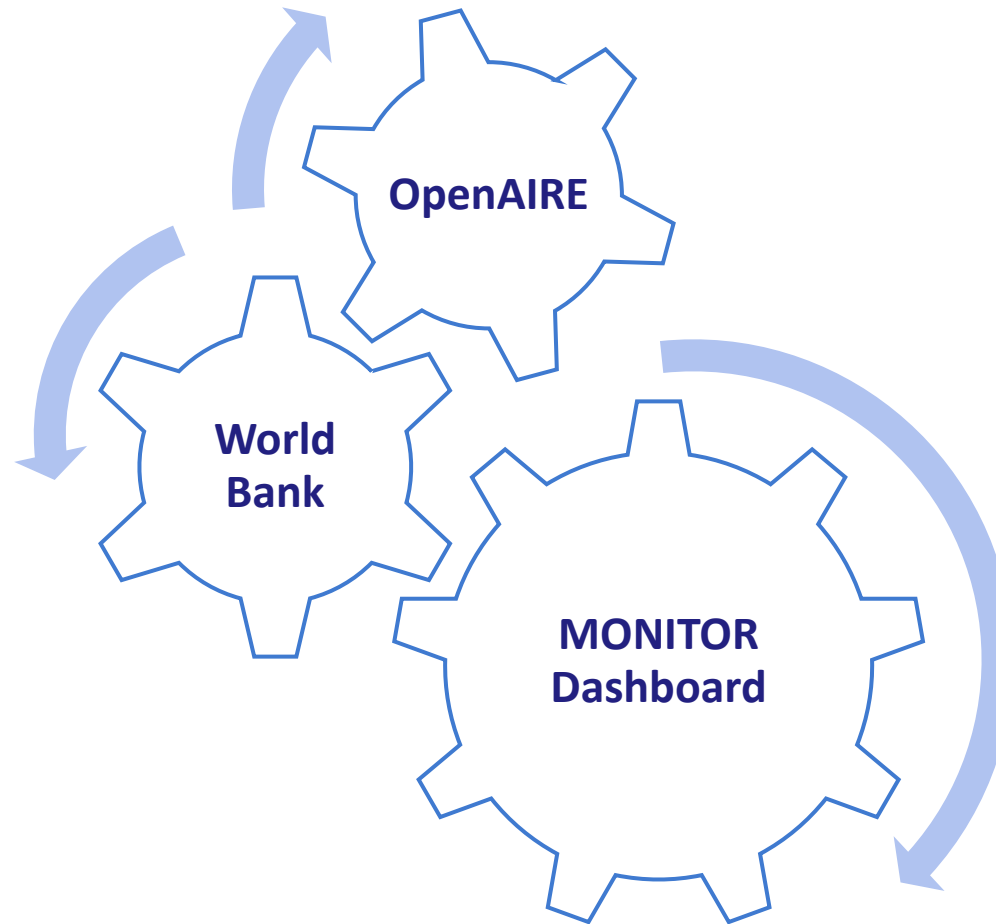


Leonidas Pispiringas

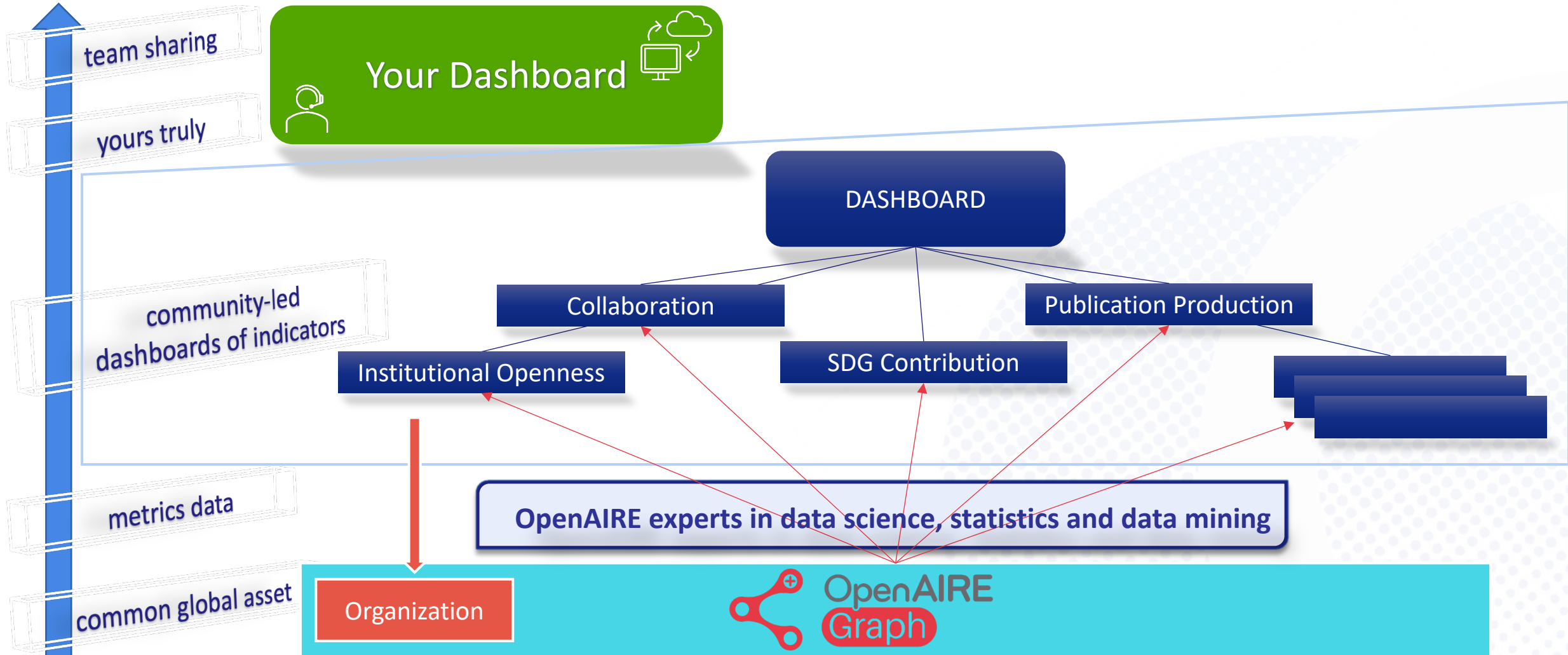
Institutional Dashboard Manager | OpenAIRE



# Iteration - Collaboration



# The Big Picture

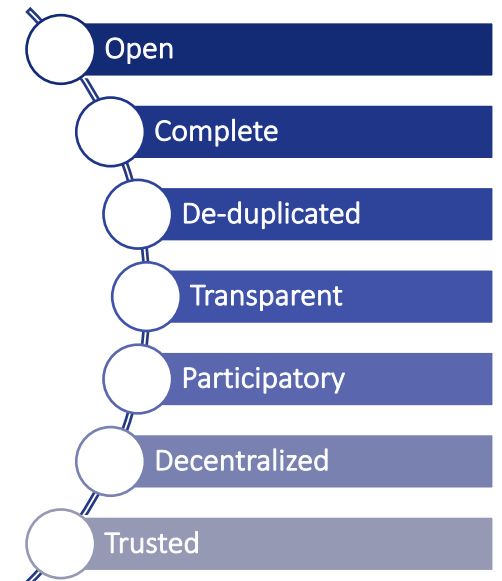


# What is OpenAIRE Graph?

- A collection of metadata describing:
  - Objects in the research lifecycle
  - Relationships among them
- Collected from ~70.000 trusted sources
- Further enriched with metadata and links provided by:
  - End-users providing links from scientific products to projects, funders, communities, or other products
  - Full-text mining algorithms over ~10Mi Open Access article full-texts
  - Research infrastructure scholarly services, bridged to the graph via OpenAIRE

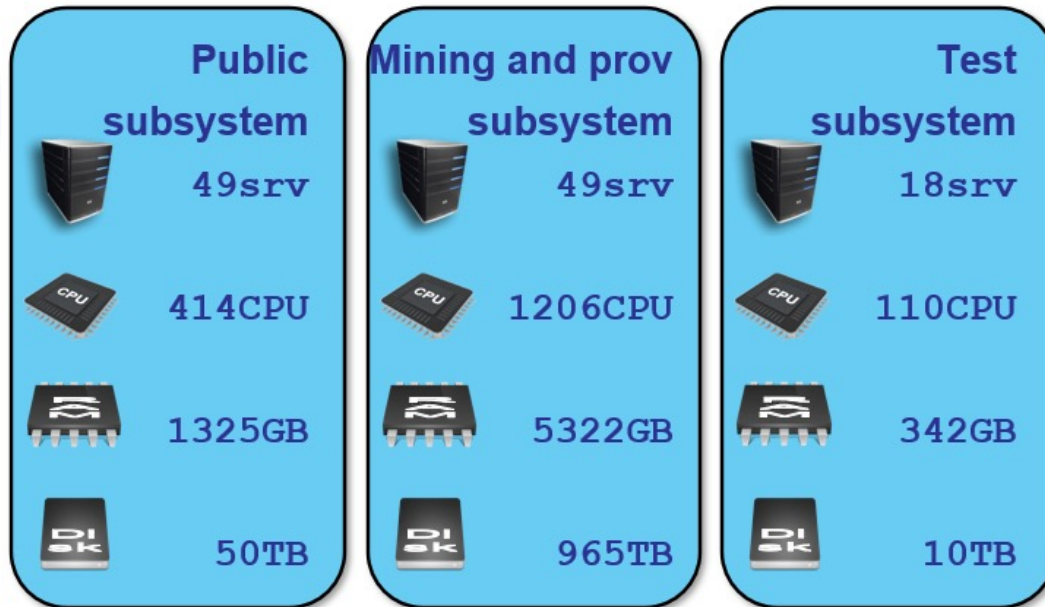


OpenAIRE  
Graph

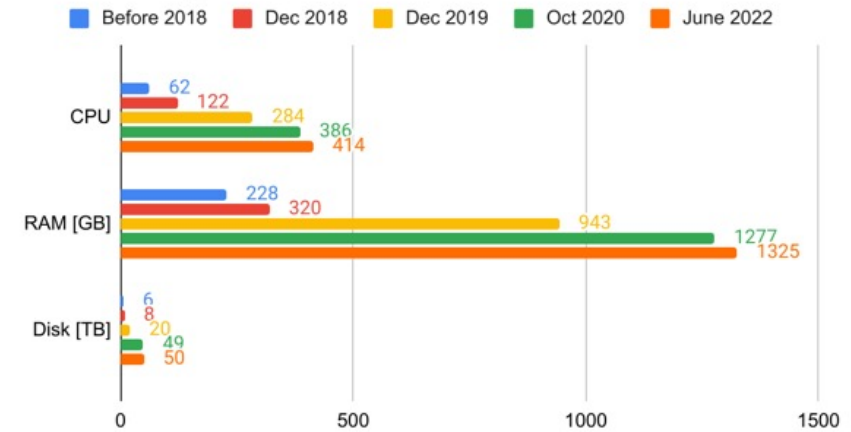




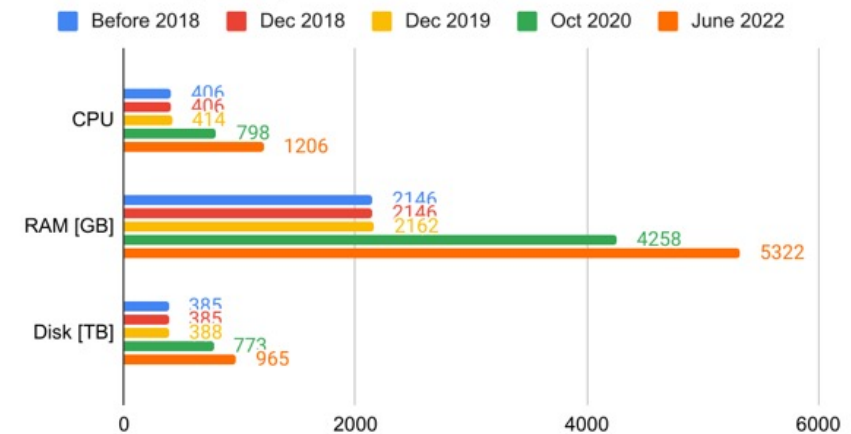
## System Infrastructure



Public subsystem resources



Data mining and provision subsystem resources



## Data quality. What can go wrong? (1)

- Identified usually in MONITOR in a few cases
- Relatively small numbers of research products
  - Disambiguation of the institutions' different names
  - Issue with the repository's metadata records during the aggregation (e.g. missing mandatory metadata field of the research product type)
  - Registration of the institution's repository(ies) sets, validates and enriches the affiliation information
  - Datasets affiliation comes from the institution's data sources

## Data quality. What can go wrong? (2)

- Identified usually in MONITOR in a few cases
- No funding information (Projects, EC projects)
  - Funders must interact with OpenAIRE and provide information for their funded projects: <https://monitor.openaire.eu/about/how-it-works>
  - Disambiguation issue in OpenOrgs



# Data quality. How to improve?

- Data sources before entering OpenAIRE Graph
  - PROVIDE Dashboard
  - Validation & Registration of your data sources
  - Compatibility with the OpenAIRE Guidelines
- Data sources that are part of the OpenAIRE graph
  - OpenOrgs Disambiguation
  - Improve the compatibility level with the OpenAIRE Guidelines

# What are my Data Sources?

- Institutional Repositories
  - Literature
  - Data
  - Software
- CRIS (Current Research Information Systems)
- Journals

# How can I register my data sources?

- Through the PROVIDE Dashboard service
- Prerequisites:
  - Metadata records:
    - Exposed via the OAI-PMH protocol
    - Compatible with the OpenAIRE Guidelines

# Why should I register my data sources?

- **EOSC:** OpenAIRE is your entry/gateway
  - If you are compliant with the OpenAIRE guidelines then
  - You will be onboarding the EOSC Portal Catalogue and Marketplace integrated platform
- **Improved interoperability** to meet latest IT and repository standards
  - Your content is more:
    - Contextualised (links and relationships to other types of research outcomes and entities)
    - Flexible (use of different and improved vocabularies)
    - Embedded in the R&I ecosystem (align with Open Science mandates and standards and support for well-established metadata schemas)
- This is your road to **FAIRness**
  - If you are compliant with the OpenAIRE guidelines then
  - you are also FAIR enough
- **Accurate and qualitative metrics**

# OpenAIRE Guidelines

- A common metadata framework for exchanging minimum metadata information
- Provide orientation to define and implement your local data management policies according to the requirements of OpenAIRE
- Expose publications, datasets and CRIS metadata via the OAI-PMH protocol in order to integrate with OpenAIRE infrastructure

# OpenOrgs

- Information ambiguity
  - Information about the organizations is distributed in numerous data sources
  - Organizations may appear with different names (legal names, acronyms, short names, alternatives...)
  - The organization's structure may not be so clear: faculties, departments, branches, detachments etc.
  
- Information disambiguation
  - **de-duplicate organizations**
  - **identify parent-child relationship**



# OpenOrgs activity pillars

- Automatic suggestion of duplicates
  - periodically produce new suggestions for the curators when a new source of organization objects is registered in OpenAIRE
- Metadata curation
  - Curators can enrich the metadata description of the organization entities thus improving the discoverability / completeness
- Management of duplicates
  - Task that only humans can carry out precisely. Curators manage the ambiguities in the data lack of the organization mentions aggregated by OpenAIRE

# We are here to assist you!

- PROVIDE Dashboard Validation & Registration guide:
  - <https://www.openaire.eu/validator-registration-guide>
- Support
  - [helpdesk@openaire.eu](mailto:helpdesk@openaire.eu)
  - Dedicated calls with our experts



**THE WORLD BANK**

# MONITOR DASHBOARD

Institutional needs, requirements and  
the MONITOR experience

---

Katie Bannon

Sr Information Officer | World Bank Group



# World Bank & the OpenAIRE Monitor Dashboard



## Why the OpenAIRE Monitor Dashboard?

- WBG is currently trying to understand how we are doing compared to Open Science standards
- Reliable data is essential to persuading management
- WBG's Data Roadmap has an indicator: 'existence of a public Open Science Dashboard'
- The World Bank issues an Annual Report on the Access to Info Policy



## OpenAIRE vs building in-house?

- Cost-effective to leverage technical & intellectual investment by OpenAire and not rebuild
- Transparent methodology makes dashboard defensible within the institution
- Self-monitoring inherently biased
- Third-party monitoring enforces standards and inter-operability



## Comparable indicators encourage competition

- Institutions tend to respond to comparisons to their peers
- International Aid Transparency Initiative's [Publish What You Fund](#) incentivizes Access to Information work
- Good tool to encourage collaboration amongst institutions

*"Development knowledge is part of the 'global commons': it belongs to everyone, and everyone should benefit from it."*

James Wolfensohn, former World Bank President, [speech](#) on 10/1/96.

*"Today, the Bank remains the largest single source of development knowledge. But knowledge must be opened to all... We need to throw open the doors, recognizing that others can find and create their own solutions. And this open research revolution is underway"*

Robert Zoellick, former World Bank President, [speech](#) on 9/29/2010.



# Experience Building the Dashboard



## Open Knowledge Repository

The OKR is the WBG's Open Access repository (on dspace) and is focused on high value publications & research with about 30K+ content. It was already following the OAI-PMH metadata standard and registered in DOAR so OKR content has been ingested into the OS Dashboard.

[Pending: data validation by WB, downloads](#)



## Documents & Reports

The WB's official disclosure mechanism for the Access to Information Policy contains 400K+ documents and was built in-house. Challenges include: the API is not aligned with OAI-PMH so needed to map, partial DOIs, some content duplication with OKR, a lot of gray literature and project records.

[Issue: defining clear inclusion criteria](#)



## Journal Articles

It is very difficult to track where researchers are submitting journal articles, though some deposit in OKR. Challenges include clearly defining WBG affiliation, measuring pre-prints vs final version, limited adoption of ORCID, expensive indexes (eg WoS), & messy metadata.

[Issue: identifying complete set articles](#)



## Datasets

The Devt Data Catalog is the umbrella catalog for WB Open Data. The catalog has an API but needs to be mapped to OAI-PMH. There is interest in exploring tools for automating FAIR assessments.

[Pending: mapping API, FAIR measurement](#)



## Research Code

WB researchers often store code in GitHub and occasionally use Zenodo. Future plans for scaling up code capture/archiving will be aligned with standards (in progress). Reached decision that MIT and BSD-3 license can be used for open research code.

[Pending: scaling up plans for open code](#)



## Other Research Products

The WBG has other public content such photo catalogs, digital archival records, and maps. This content is not a research product but rather raw research material. Should this sort of content be included?

[Issue: defining clear inclusion criteria](#)



## Impact

Researchers are eager to understand how their content is used but it is very difficult to measure usage from fragmented platforms with unintegrated, non-standard, metadata. Providing more comprehensive and insightful stats for researchers would be invaluable.

[Issue: adopting inter-operable metadata](#)

# MONITOR INDICATORS

## New & Upcoming

---

Ioanna Grypari

Indicators Lead | ATHENA Research Center





## BEHIND THE INDICATORS..

## WHY MONITOR?



### Know Thyself

Resources  
Research Output  
Open Science Uptake  
Collaborations  
Visibility  
Impact



### Understand Thyself

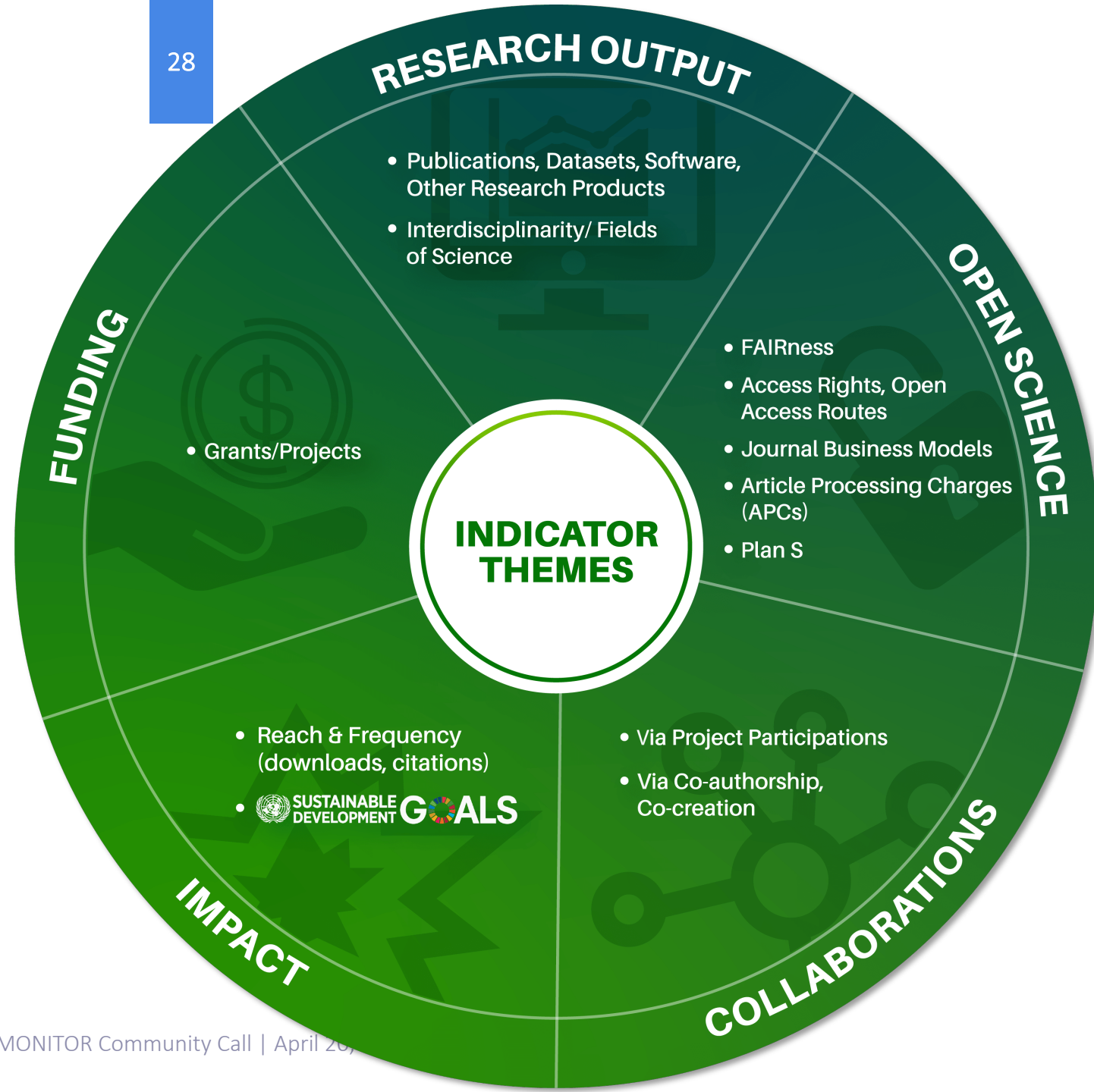
Pathways  
Insights  
Opportunities



### Position Thyself

Decision making  
Reporting  
Story telling

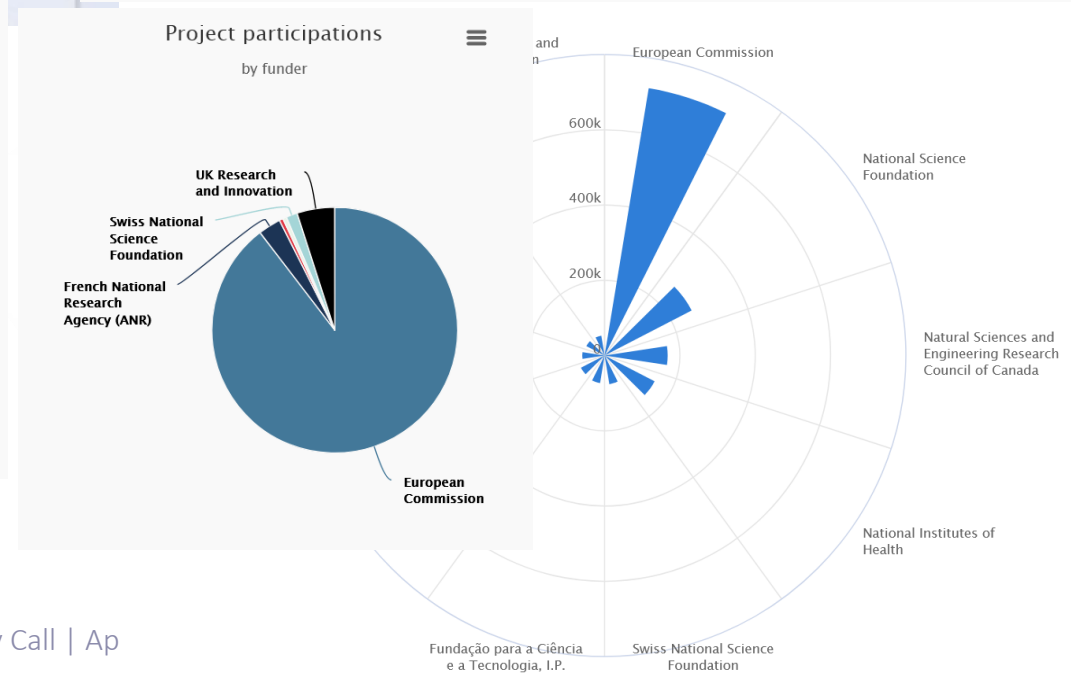
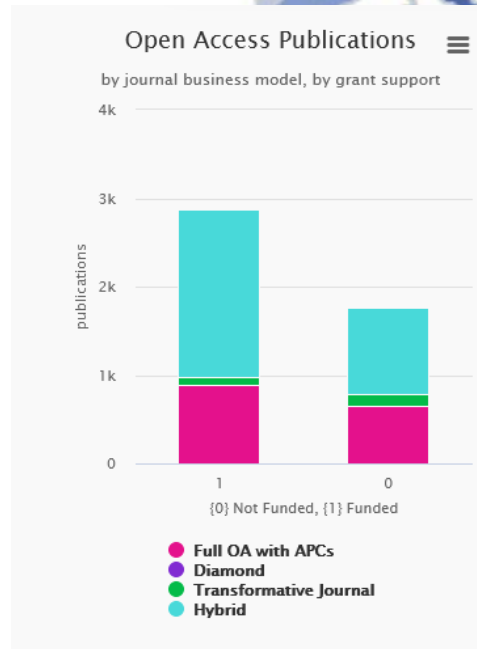
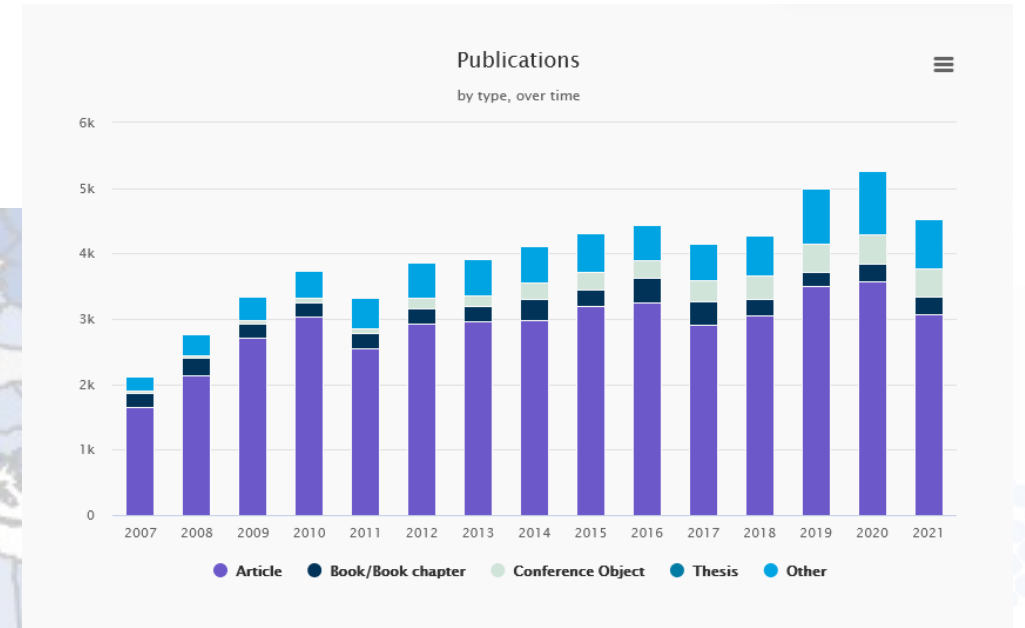
# Indicators bird's-eye view



# Indicators Breakdown

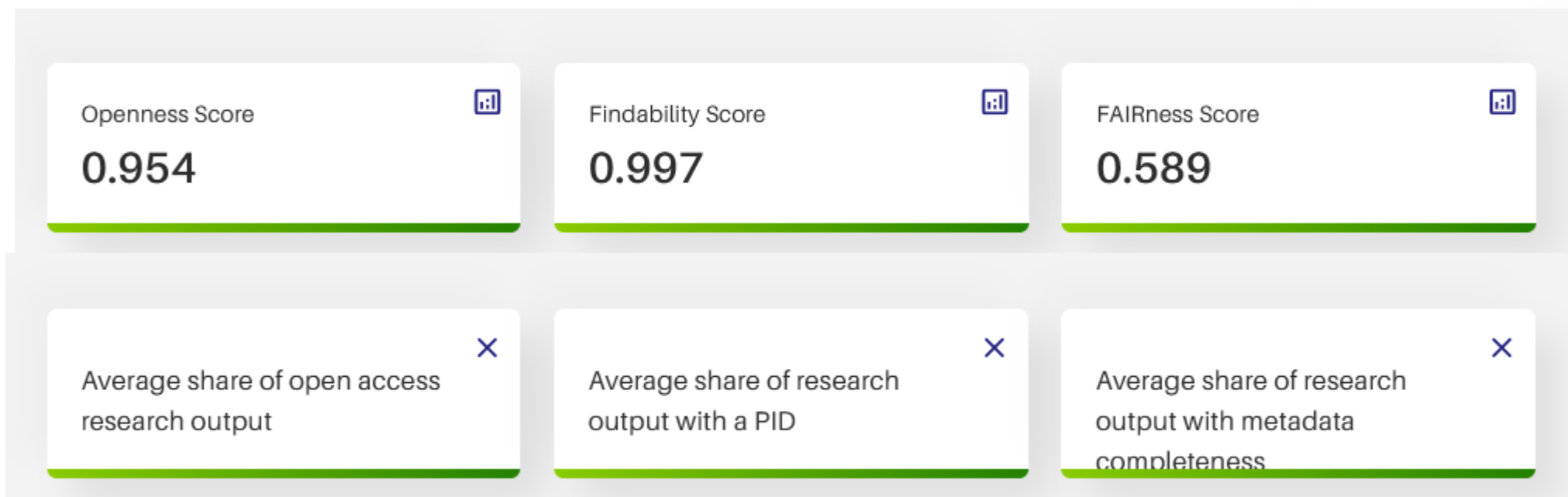
## By fields of interest

- research product type
- time
- countries
- data sources
- funders
- ...



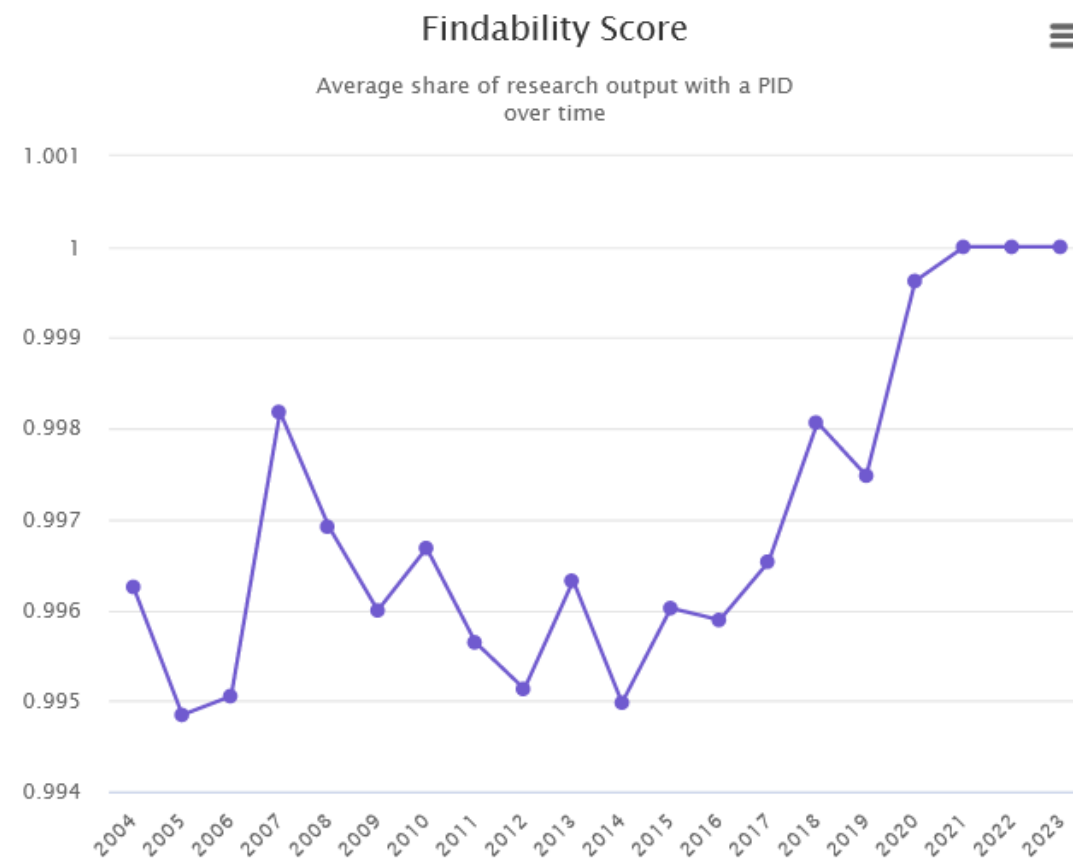
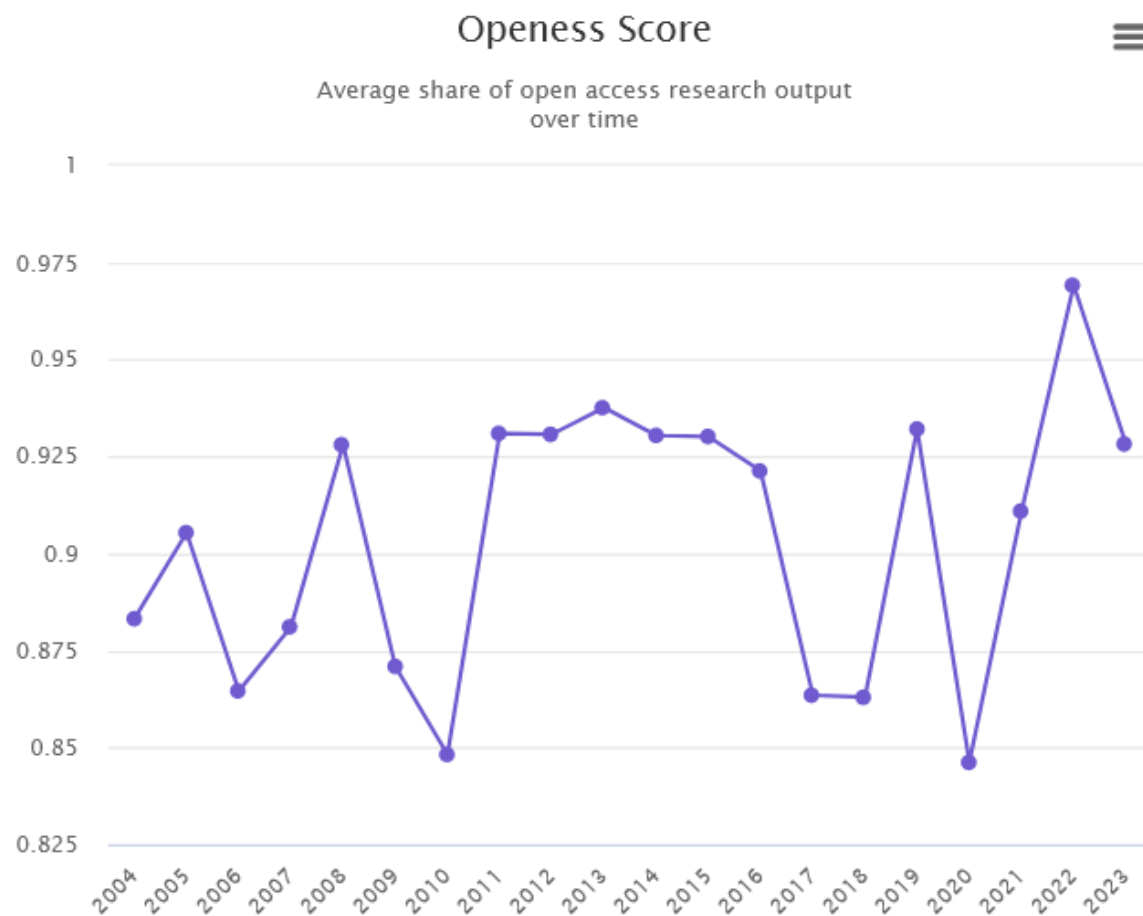
# New Indicators OS Composites

Open Science composite indicators for organizations



# New Indicators OS Composites

## Open Science composite indicators for organizations



# New Indicators UN SDGs

- We have developed a classification scheme for UN Sustainable Development Goals, to view contributions of research towards complex challenges for humanity such as climate change, biodiversity loss, pollution and poverty reduction.

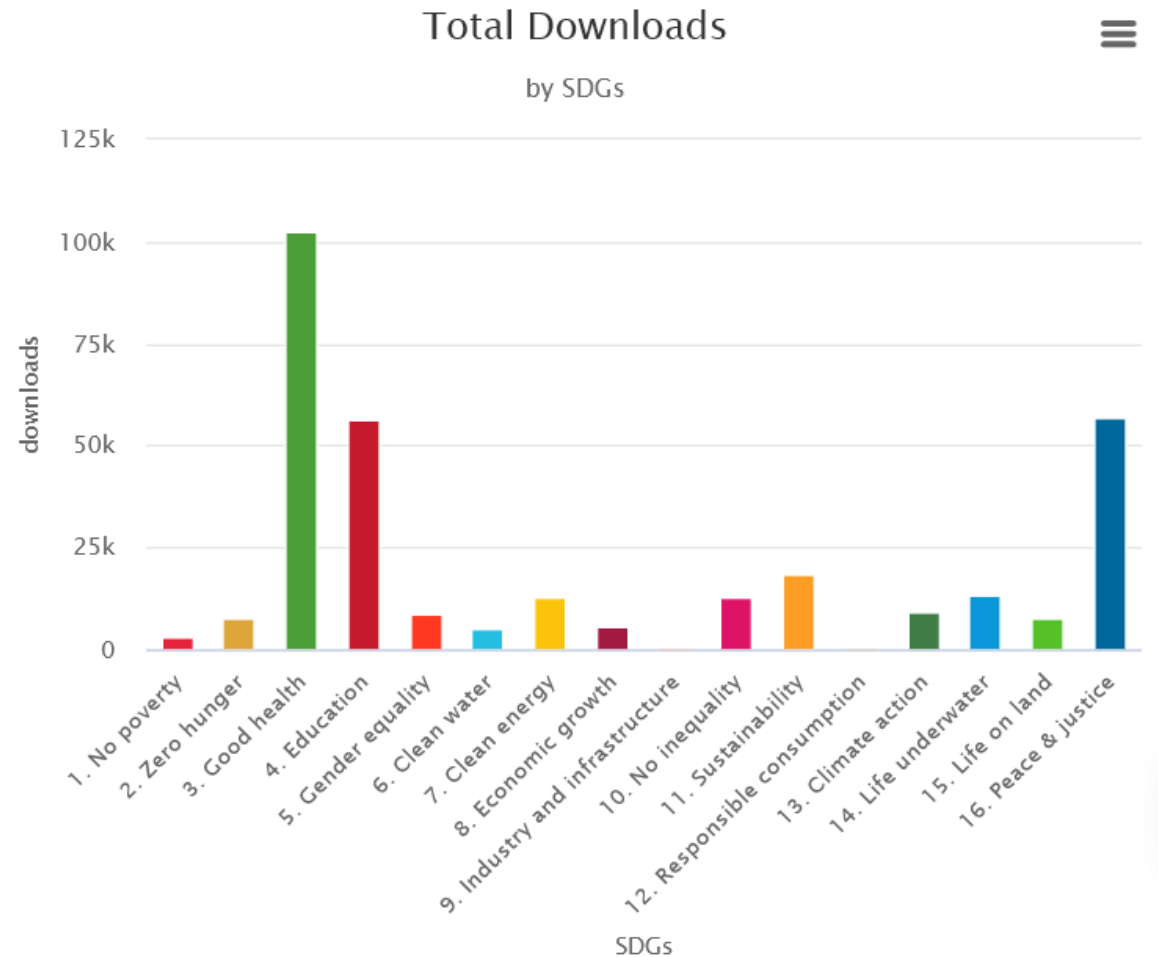
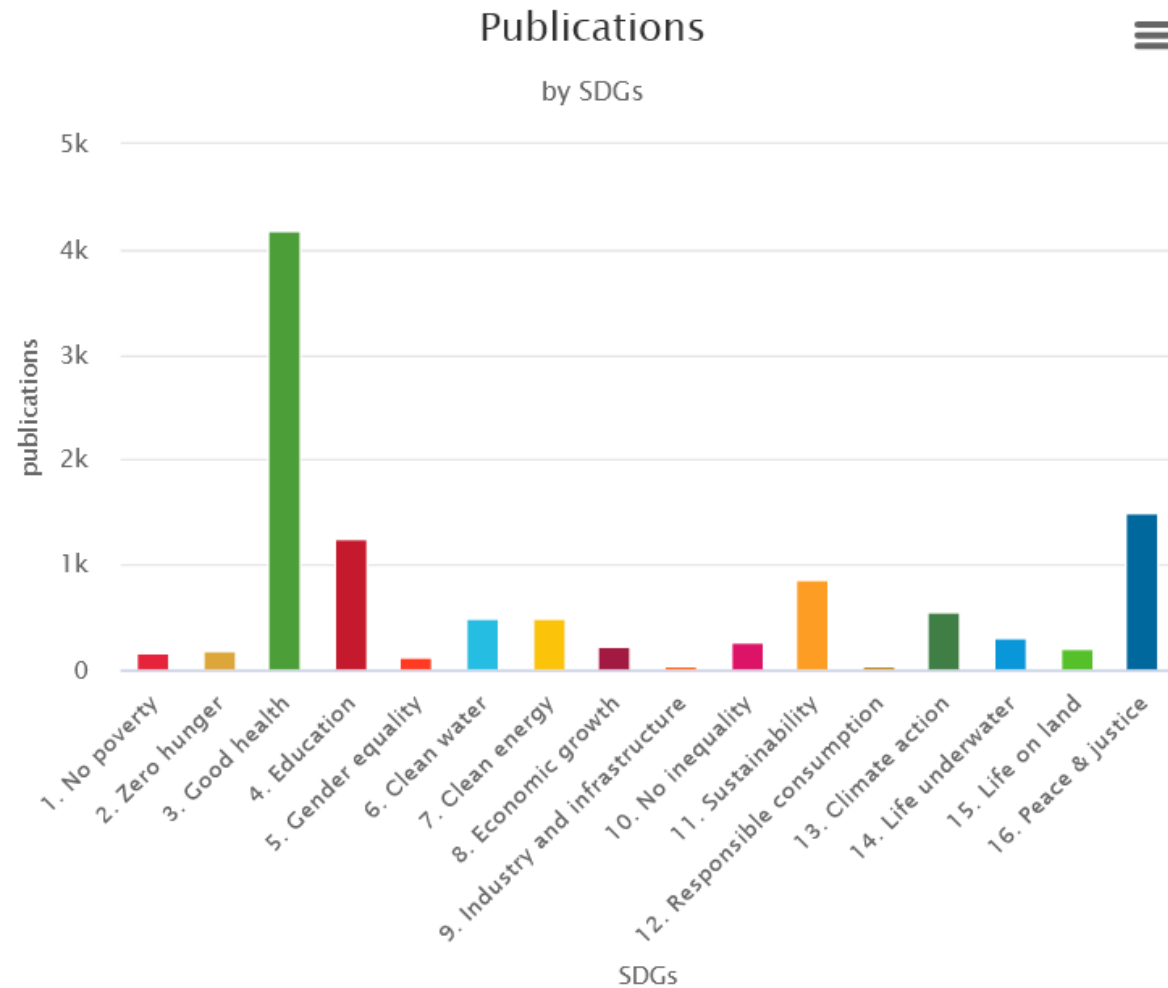
## SUSTAINABLE DEVELOPMENT GOALS



<https://explore.openaire.eu/sdgs>



# New Indicators UN SDGs



# Additional Dimension of Analysis

## Fields of Science (FoS) Classification Levels 1, 2, 3

**01 Natural Sciences**

02 Engineering And Technology

03 Medical And Health Sciences

04 Agricultural And Veterinary Sciences

05 Social Sciences

06 Humanities And The Arts

### 01 natural sciences

**0101 mathematics**

010101 Applied Mathematics

010102 General Mathematics

010103 Numerical & Computational Mathematics

010104 Statistics & Probability

**0102 computer and information sciences**

010201 Computation Theory & Mathematics

**0103 physical sciences**

010301 Acoustics

010302 Applied Physics

010303 Astronomy & Astrophysics

010304 Chemical Physics

010305 Fluids & Plasmas

010306 General Physics

010307 Mathematical Physics

010308 Nuclear & Particles Physics

010309 Optics

**0104 chemical sciences**

010401 Analytical Chemistry

010402 General Chemistry

010403 Inorganic & Nuclear Chemistry

010404 Medicinal & Biomolecular

**0105 earth and related environmental sciences**

010501 Environmental Sciences

010502 Geochemistry & Geophysics

010503 Geology

**0106 biological sciences**

010601 Ecology

010602 Entomology

010603 Evolutionary Biology

010604 Marine Biology

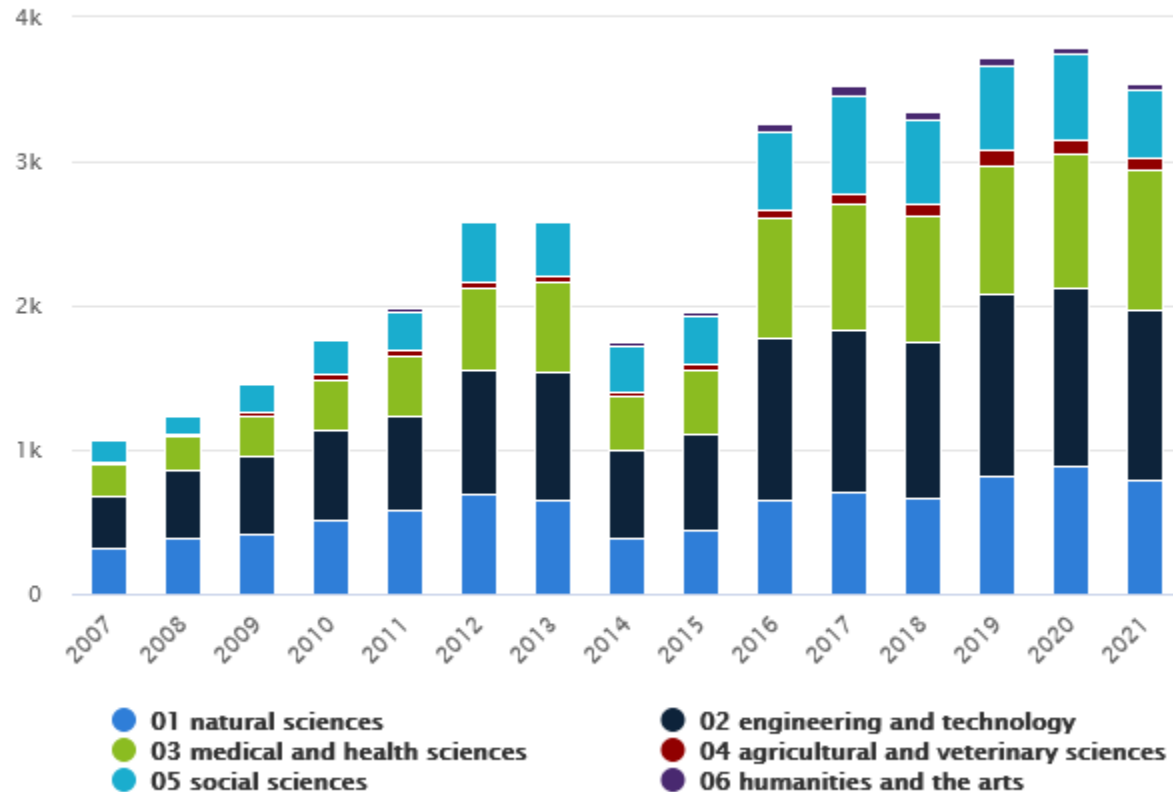
Hydrobiology

HELP

# Indicators Breakdown by FoS

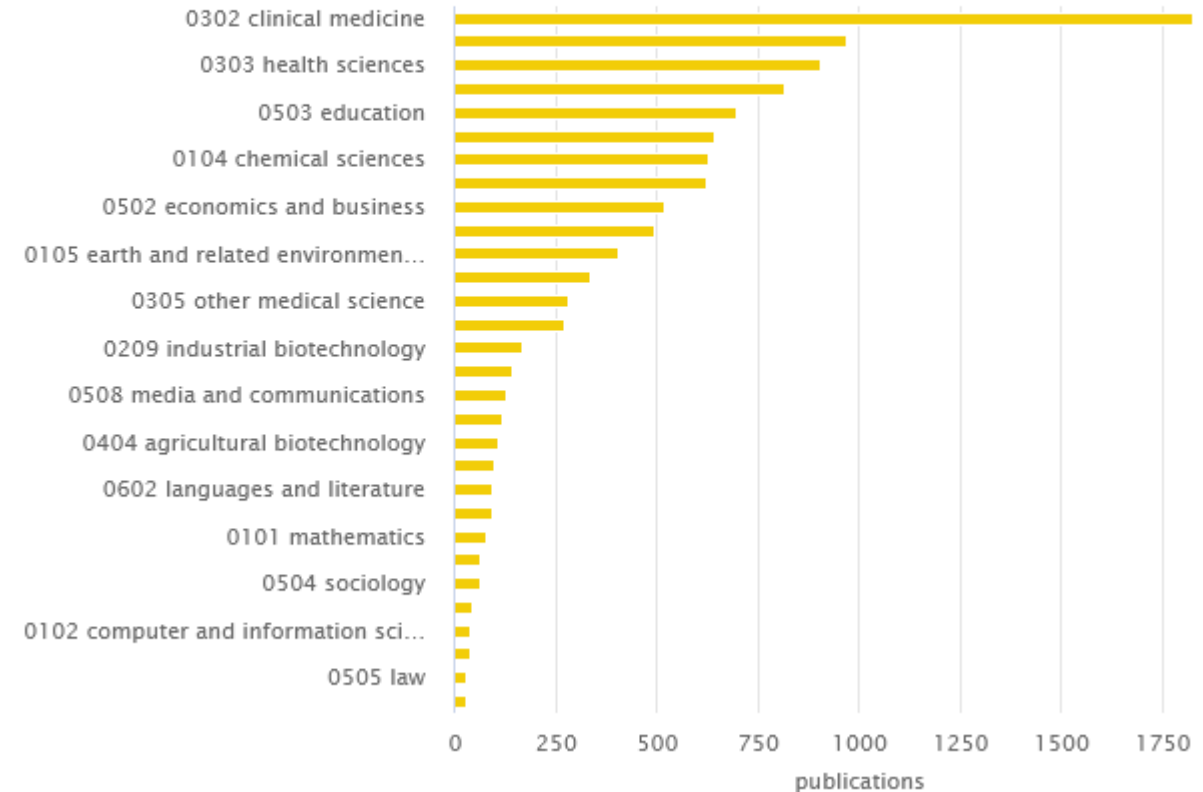
## Publications

by FoS (level 1), over time

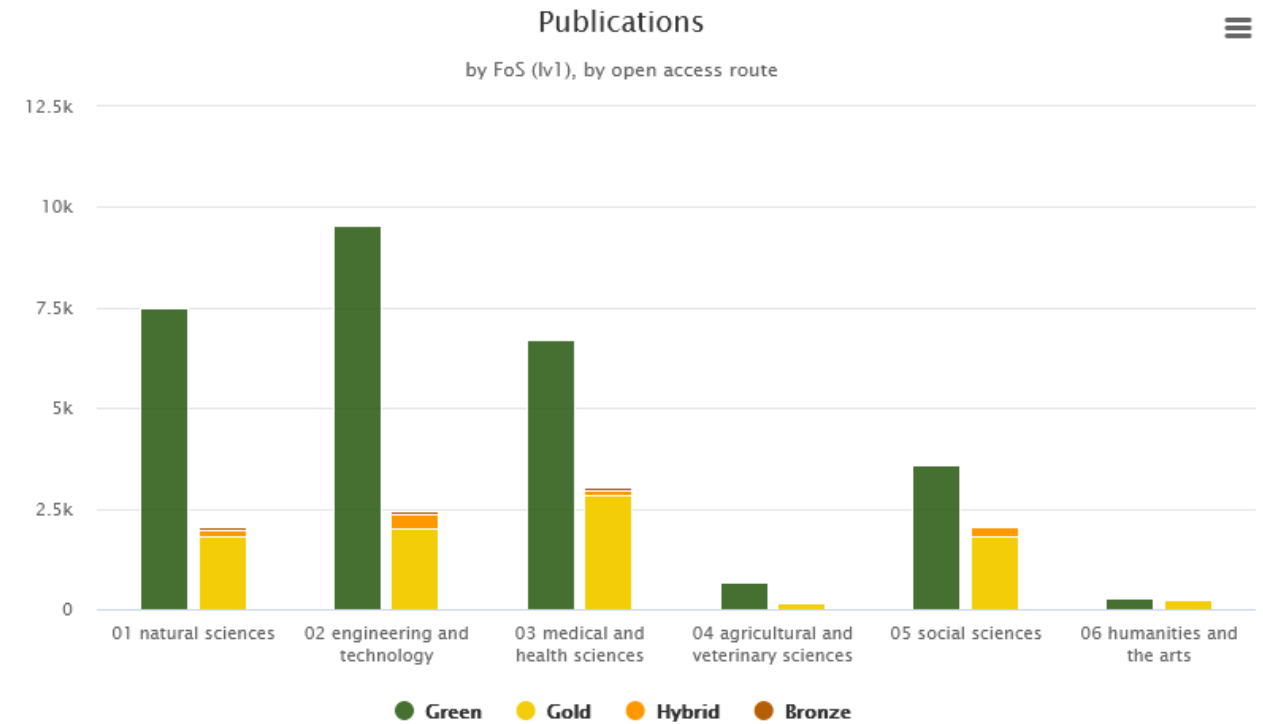
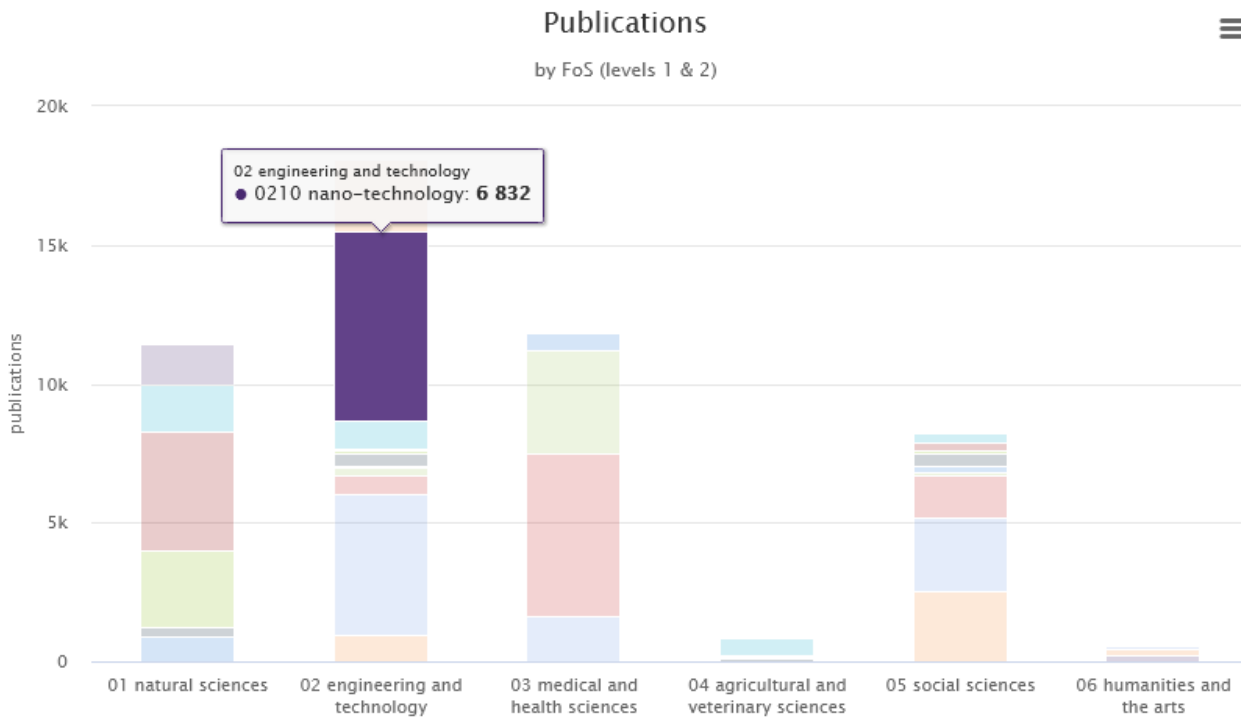


## Most open FoS (level 2)

by number of gold OA publications



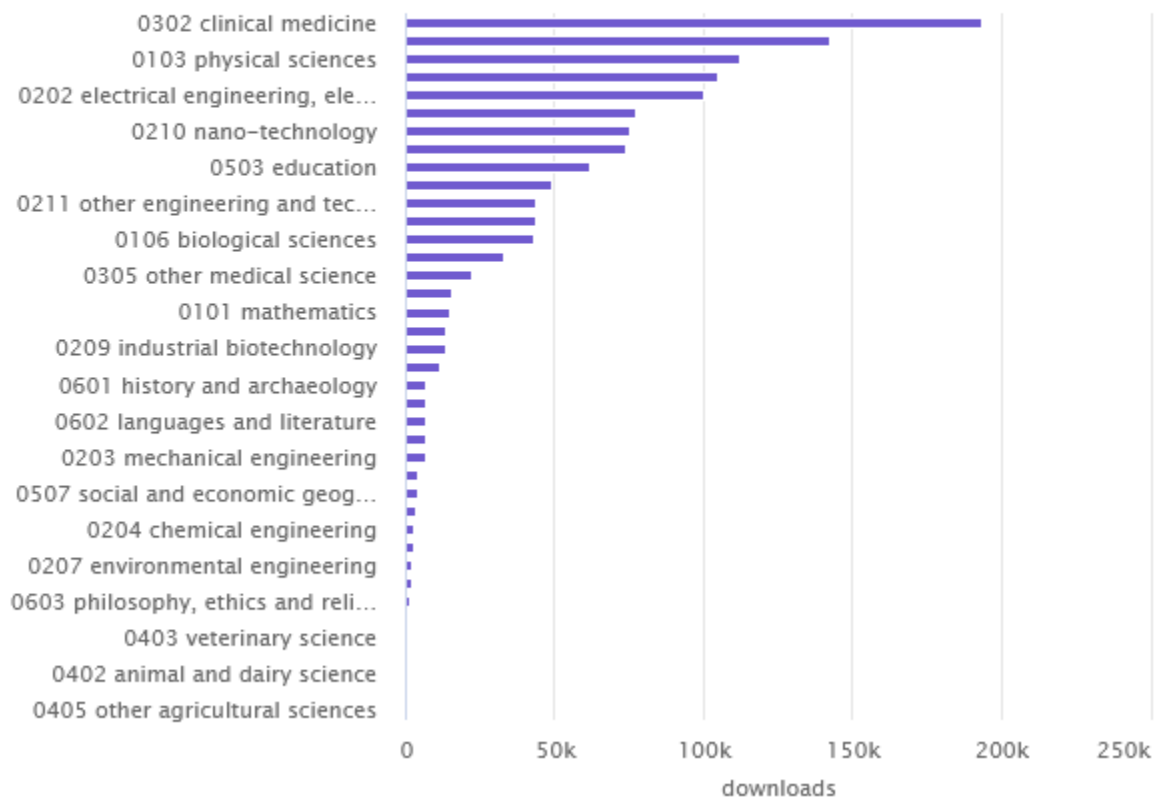
# Indicators Breakdown by FoS



# Indicators Breakdown by FoS

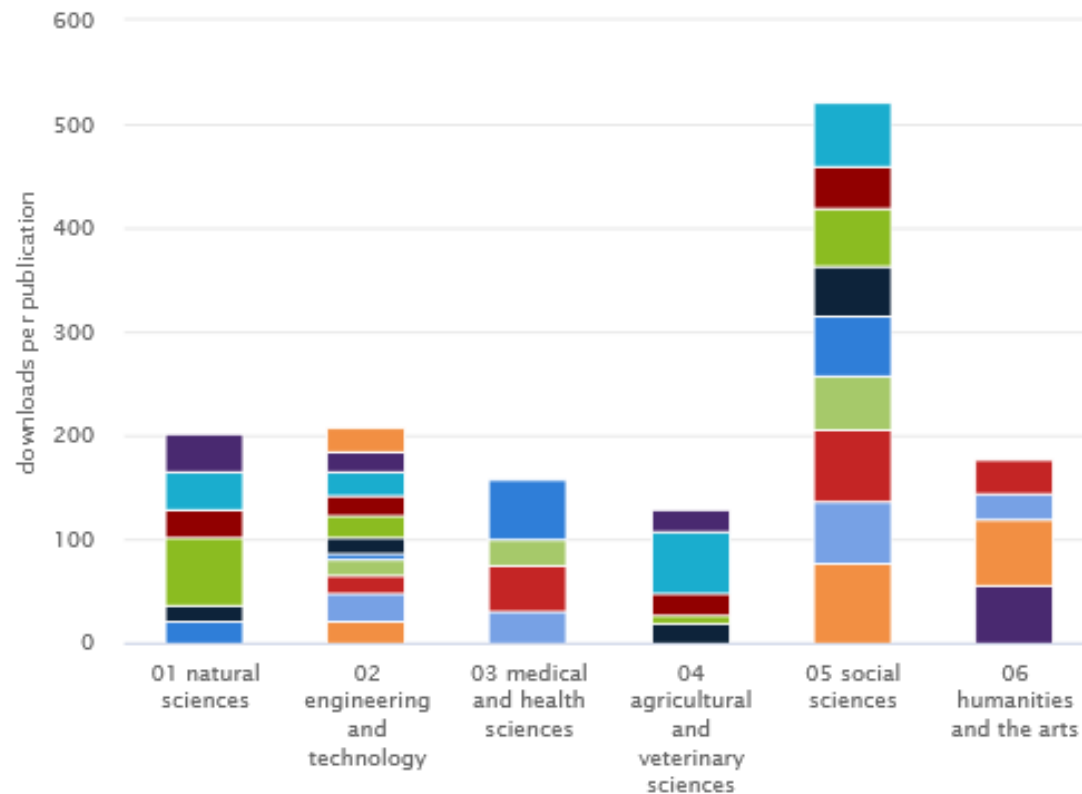
## Most downloaded FoS (level 2)

by number of publication downloads



## Average Downloads per Publication

by FoS (levels 1 & 2)



# Upcoming Indicators

## Citations

- Citation count, “incubation” citation count, RAM, page rank  
(<https://graph.openaire.eu/docs/graph-production-workflow/indicators-ingestion/impact-indicators>)

## Article Processing Charges

- OpenAPC

## Horizon Europe projects indicators

### Plan S Compliance

- PlanS funder grant-supported output
- Transformative agreements

### Analysis/Breakdown by

- **Multi-disciplinarity**

**Great! When can I see these on my dashboard?**

**New: May 2023**

**Upcoming: September 2023**

**How do I get started with a new dashboard?**

**Here:**  
**[monitor.openaire.eu/get-started](https://monitor.openaire.eu/get-started)**



# THANKS

---

**Email**

[igrypari@athenarc.gr](mailto:igrypari@athenarc.gr)

[leonidas.pispiringas@openaire.eu](mailto:leonidas.pispiringas@openaire.eu)

---

**Twitter**

[@openaire\\_eu](https://twitter.com/openaire_eu)

---