Dimensions

Dimensions as scientific data infrastructure: funding data, international flows, and open access

Michele Pasin | VP Data Products | Digital Science

CLACSO / Tracking Research Funding Flows in the Global South

Buenos Aires, April 2026

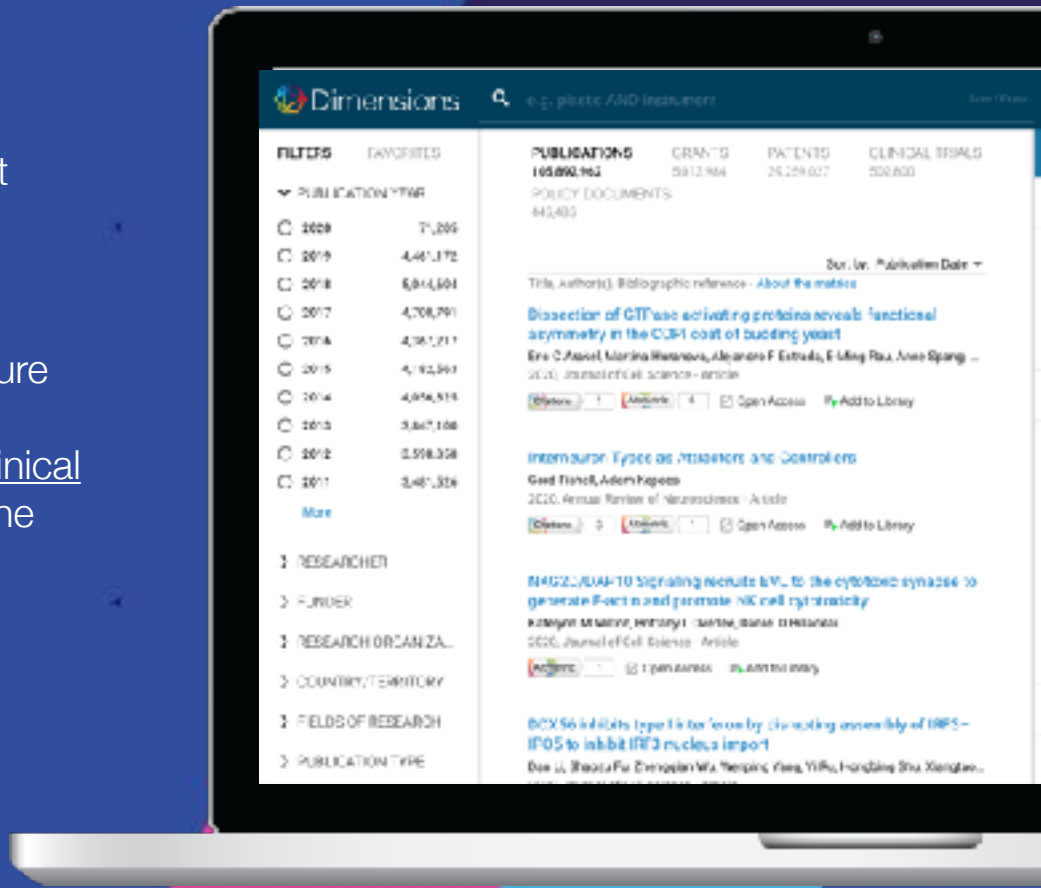
Part of DIGITAL science

What is Dimensions?

Dimensions delivers an array of search and discovery, analytical, and research management tools, all in a single platform.

- **Moving beyond citations** to better capture the research lifecycle: the **links among** publications, grants, policy documents, clinical trials, datasets, patents, researchers are the key to understanding this landscape.

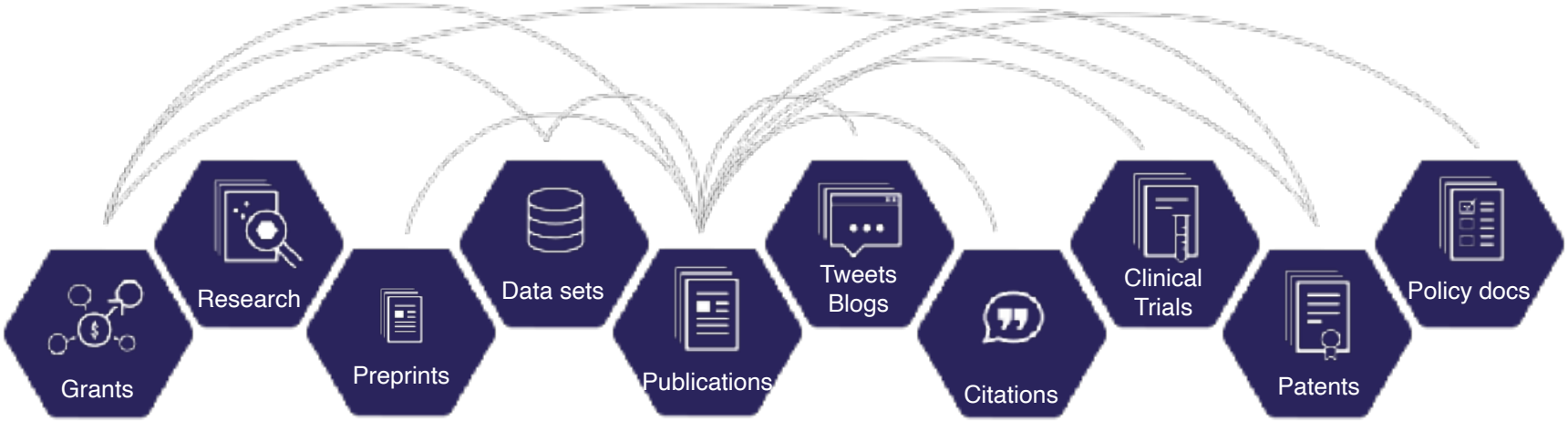
https://www.dimensions.ai



Dimensions

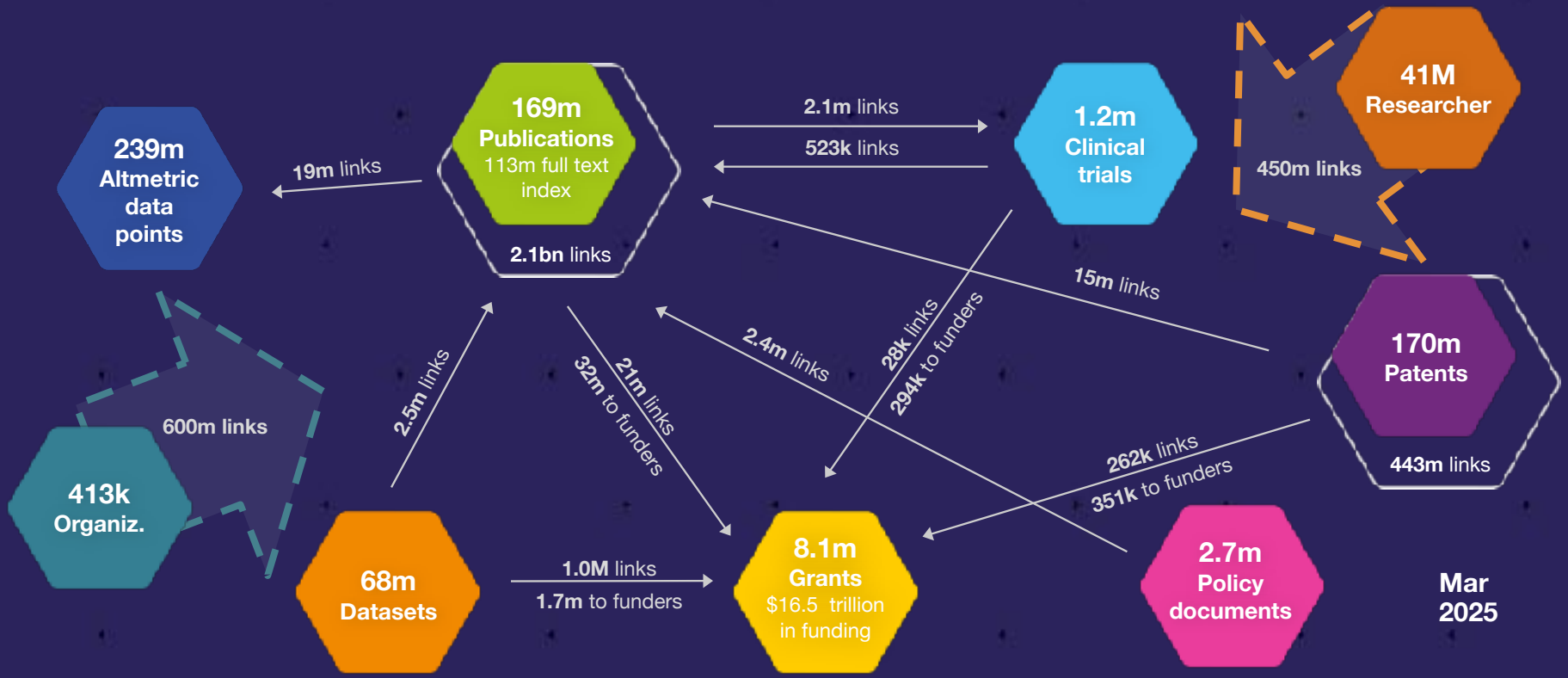
Part of DIGITAL science

We capture a much broader picture of the research landscape





The science data jungle



Dimensions

Altmetric

ifi CLAIMS

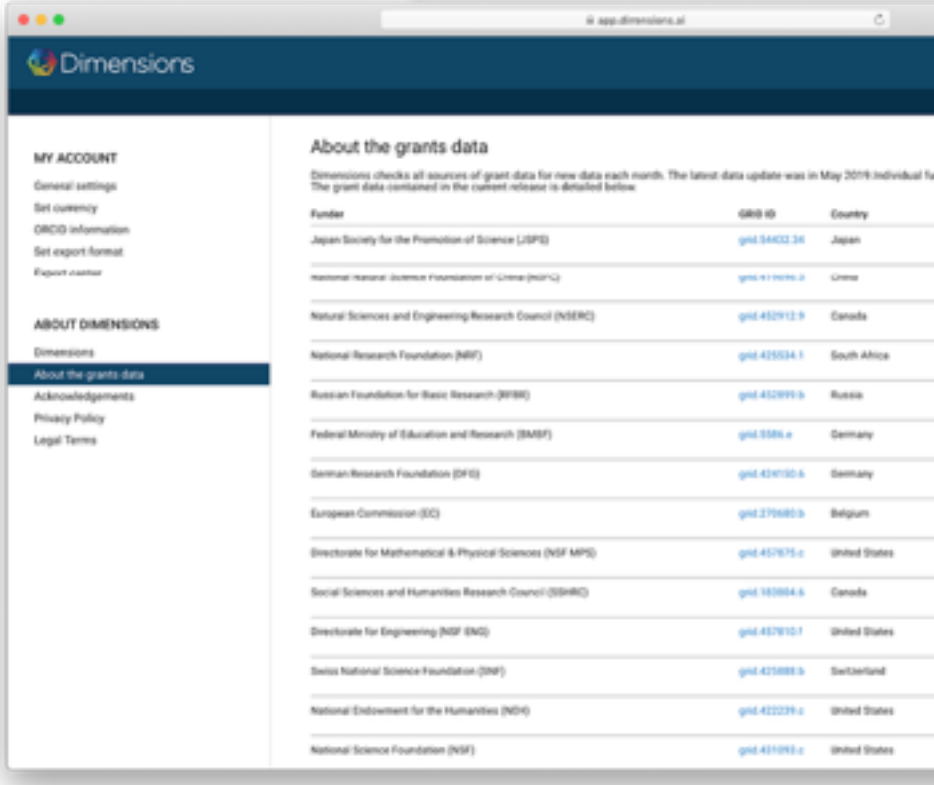
The funding data challenge

Why funding data is hard:

- Funders don't report to a single registry
- Grant amounts are often not public
- Poorly structured data endpoints (PDFs, unstable servers) & restrictive copyright terms
- Southern funders are underrepresented in global registries

Why it matters:

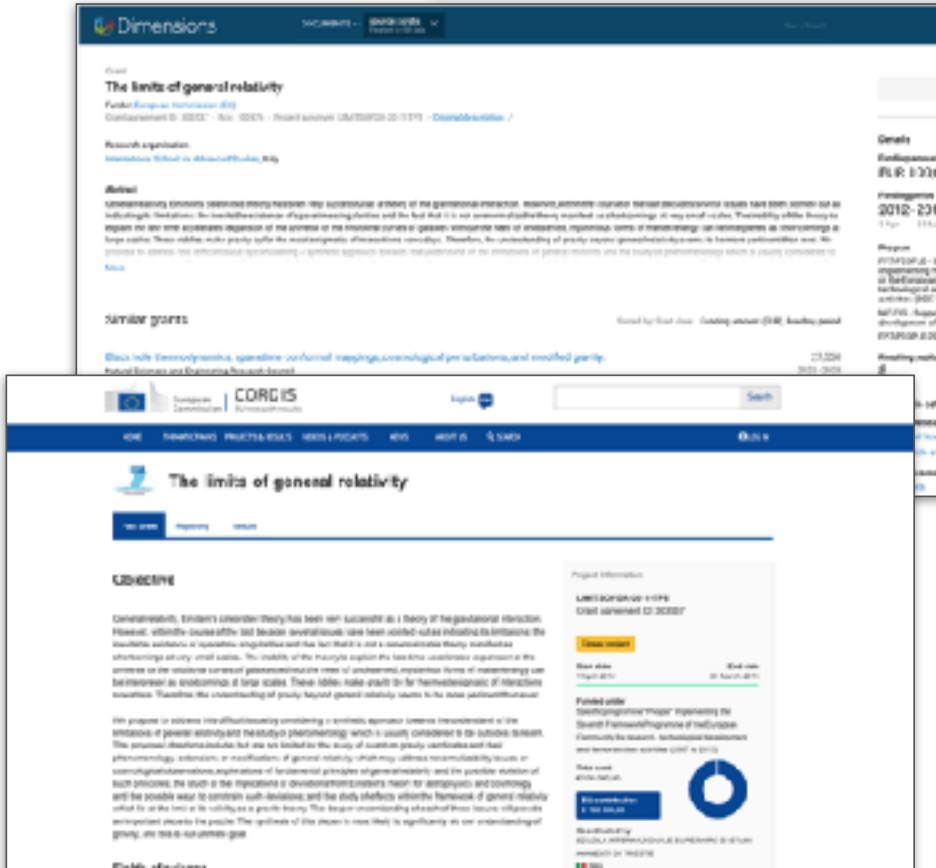
- Without good data, granting councils can't see who else is funding in their country
- Policy decisions on R&D priorities are made on incomplete pictures
- The information asymmetry favours funders over recipients



Part of DIGITAL science

Data stream 1: Grant records

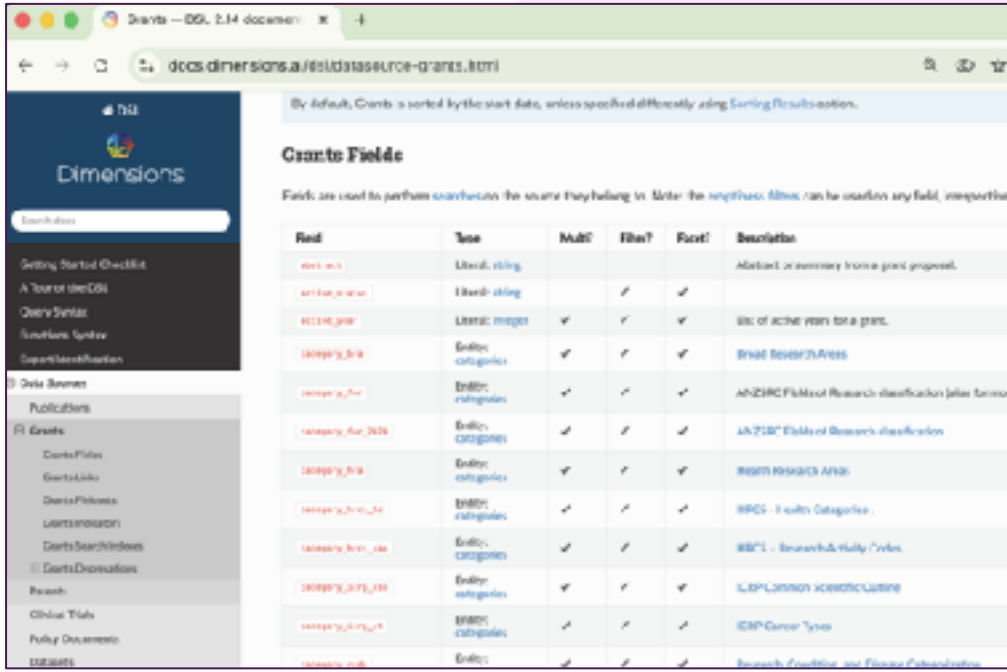
- Project funding
  - **8.1M** grants **738** funders globally
  - **$16.5** trillion of funding
- Sourcing
  - Direct relationships with funders
  - Data available via APIs & websites which we crawl (e.g. ORCID)
- Grants are excluded when they are:
  - Non-scientific research related
  - From a copyrighted website
  - Missing mandatory fields (ID, Title, Funder organization name, Funder city, Funder country
  - Removed from source website
  - Duplicates



Dimensions

Data stream 1: Grant records => Fields

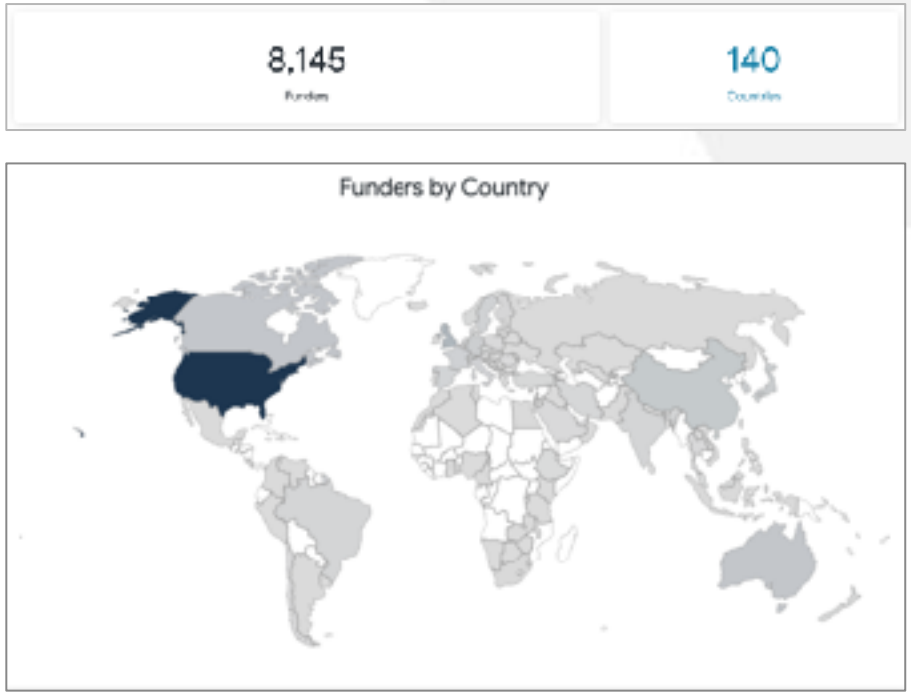
- Grants fields
  - Dimensions ID and Original ID
  - Title, Abstract, Funding Schemes, Keywords
  - Funding Period (ideally with precise Start and End Date) & Funding Amount (possibly in currencies other than USD)
  - Funding Organization
  - Researchers (Principal Investigators and Co-Principal Investigators)
  - Research Organisations (possibly affiliated with the researchers)
  - Classifications (FoR, SDGs, HRAs etc..)



The Dimensions Search Language. Data Sources: Grants https://docs.dimensions.ai/dsl/datasource-grants.html

Data stream 2: Funding acknowledgements

- Source: Full text of research documents, extracted via NLP
  - **8145** funders globally (Vs 738 with grants)
  - Publications, Patents, Clinical Trials, Datasets
- What it captures:
  - Funder name as mentioned in the document e.g. acknowledgement or funding section
  - Mapped to standardised funder IDs
  - Linked directly to the document
- Why it matters:
  - Captures funding that was never formally registered as a grant
  - Useful for informal, bilateral, or unreported flows
  - Limitation: Skews toward indexed, English-language sources (e.g. journals)



Access routes

Web App



Search & discovery; top analytical use cases

Dedicated UI; inbuilt visualizations; dashboards

In the browser, no specialized knowledge required

For everyone

API

DSL: optimised for full-text search and common structured search use cases

1k per query, one query every 2 secs

Frequent interactions where small chunks of data are retrieved / stored locally

For data & analytics teams

Google BigQuery

SQL: full power of structured search

No query limits & real-time performance

Large-scale analyses & direct integration with BI tools e.g. Tableau, Qlik, PowerBI

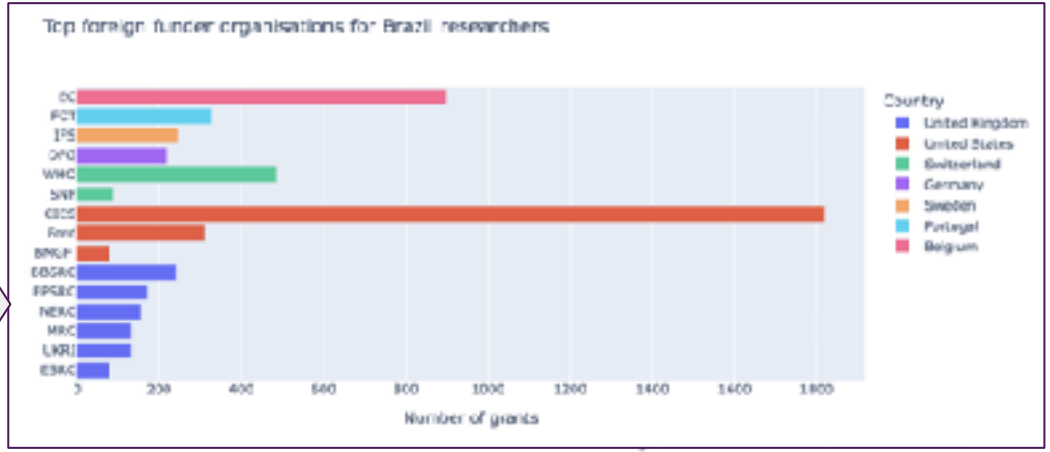
For data & analytics teams

What international flows look like in the data - Dimensions API

- **API** fields make it easy to get key statistics quickly
- Tracking internal Vs external funding using country of funder / country of recipient

- Example: Brazil
  - Grants 500k (internal) Vs 6k (external)
  - Foreign Funder_org_countries: 25
  - Belgium (EC) largest ext. funder by $$
  - UK largest ext. funder by # of Grants

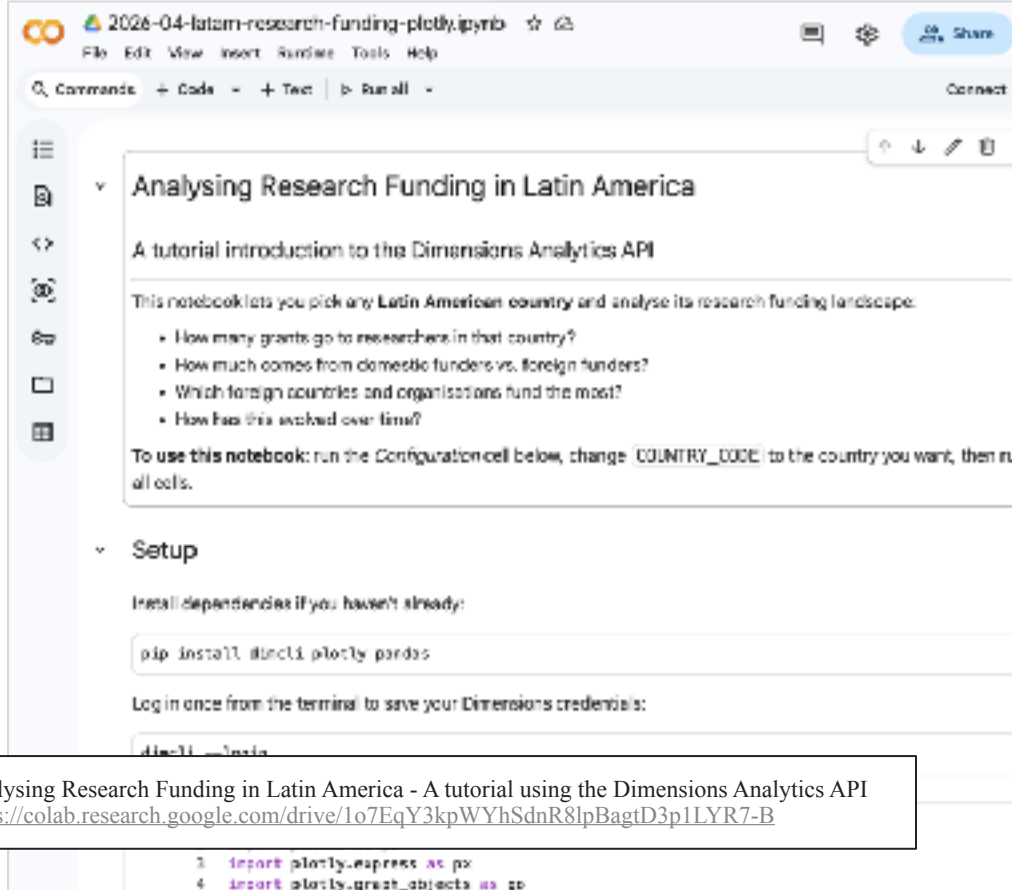
```
search grants
where research_org_countries = "BR"
  and funder_org_countries != "BR"
return funder_orgs
limit 20
```



Dimensions

What international flows look like in the data - Dimensions API

- **Colab notebook**: pick any LatAm country and generate key reports
  - Funder country breakdown
  - Grant counts: domestic vs. foreign, over time
  - Top domestic funders
  - Foreign funding amounts in USD
  - Top foreign funder organisations
  - Compare two countries side by side



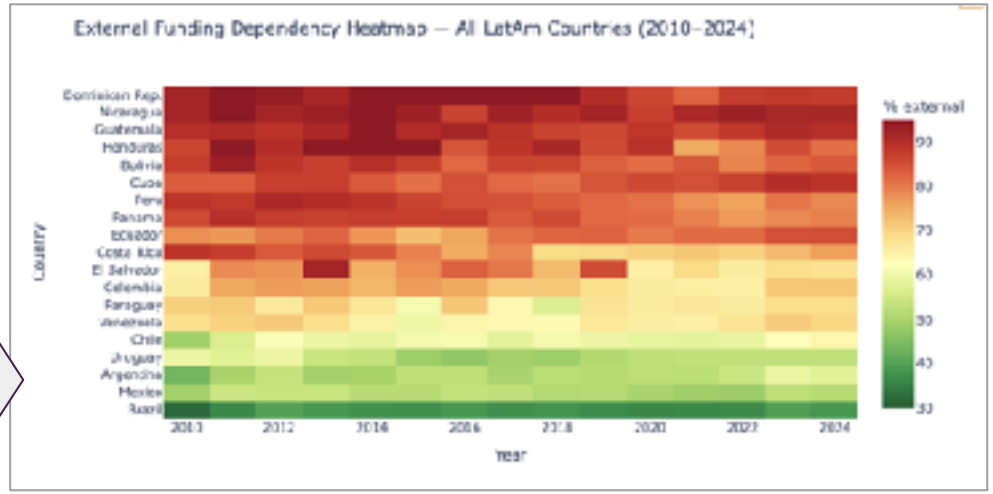
Analysing Research Funding in Latin America - A tutorial using the Dimensions Analytics API
https://colab.research.google.com/drive/1o7EqY3kpWYhSdnR8lpBagtD3p1LYR7-B

Dimensions

- BigQuery SQL supports large-scale longitudinal analysis

- Trends over time: is external dependency growing?
  - Using funded publications that acknowledge external (non-LatAm) funders
  - Reveals whether countries are building domestic funding capacity or becoming more reliant on foreign sources.

```sql
SELECT
  research_country,
  pub.year,
  SUM(CASE WHEN funder_country IN ({latam_sql})     THEN 1 ELSE 0 END) AS internal_count,
  SUM(CASE WHEN funder_country NOT IN ({latam_sql}) THEN 1 ELSE 0 END) AS external_count,
  COUNT(*)                                                             AS total_count,
  ROUND(
    SAFE_DIVIDE(
      SUM(CASE WHEN funder_country NOT IN ({latam_sql}) THEN 1 ELSE 0 END),
      COUNT(*)
    ) * 100, 2
  ) AS pct_external
FROM `dimensions-ai.data_analytics.publications` pub
CROSS JOIN UNNEST(pub.research_org_countries) AS research_country
CROSS JOIN UNNEST(pub.funder_countries)       AS funder_country
WHERE
  research_country IN ({latam_sql})
  AND pub.year BETWEEN 2010 AND 2024
  AND funder_country IS NOT NULL
GROUP BY research_country, pub.year
ORDER BY research_country, pub.year
```

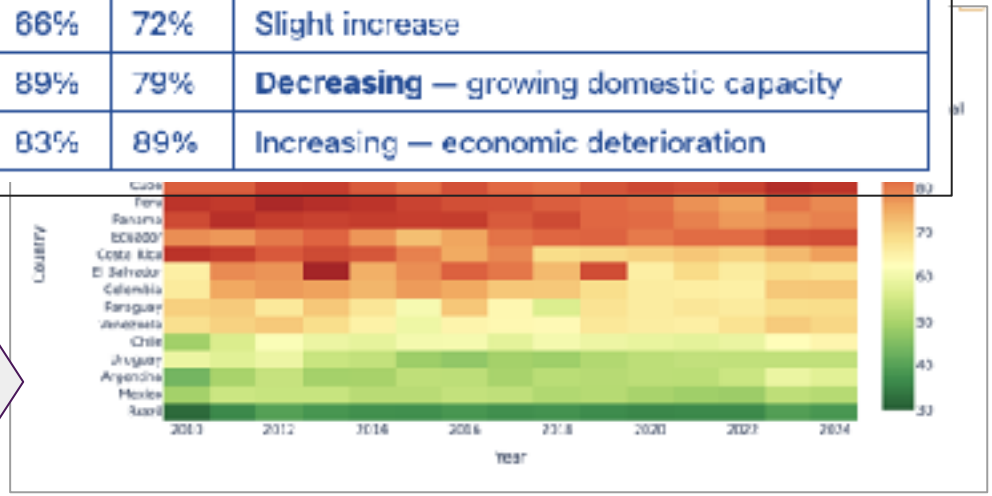


- BigQuery SQL supports large-scale longitudinal analysis

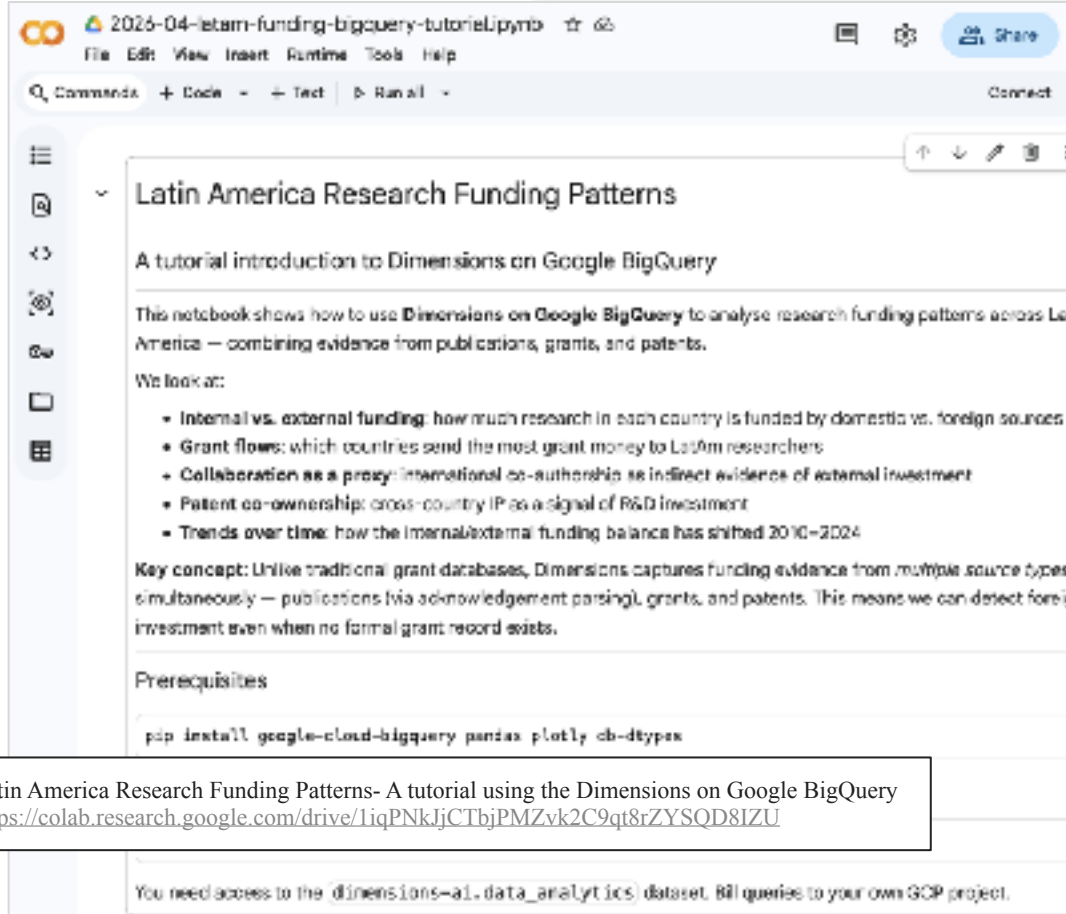
Key trend observations:

Country	2010	2024	Trend
Brazil	32%	38%	Stable, slight uptick
Mexico	49%	52%	Stable
Argentina	44%	58%	**Increasing** — CONICET underfunding
Chile	49%	64%	**Increasing** — growing international integration
Colombia	66%	72%	Slight increase
Peru	89%	79%	**Decreasing** — growing domestic capacity
Cuba	83%	89%	Increasing — economic deterioration

```sql
SELECT
  research_country,
  pub.year,
  SUM(CASE WHEN funder_country IN ({latam_sql}) THEN 1 ELSE 0 END)
  SUM(CASE WHEN funder_country NOT IN ({latam_sql}) THEN 1 ELSE 0 END)
  COUNT(*)
  ROUND(
    SAFE_DIVIDE(
      SUM(CASE WHEN funder_country NOT IN ({latam_sql}) THEN 1 ELSE 0 END),
      COUNT(*)
    ) * 100, 2
  ) AS pct_external
FROM `dimensions-ai.data_analytics.publications` pub
  CROSS JOIN UNNEST(pub.research_org_countries) AS research_country
  CROSS JOIN UNNEST(pub.funder_countries)       AS funder_country
WHERE
  research_country IN ({latam_sql})
  AND pub.year BETWEEN 2010 AND 2024
  AND funder_country IS NOT NULL
GROUP BY research_country, pub.year
ORDER BY research_country, pub.year
```



- **Colab notebook**: exploring Latin America research funding patterns
  - Publication-level funding — internal vs. external
  - Grant flows with funding amounts
  - International collaboration as a funding proxy
  - Patent cross-country co-ownership
  - Top external funders ranking
  - Trends over time — is external dependency growing?

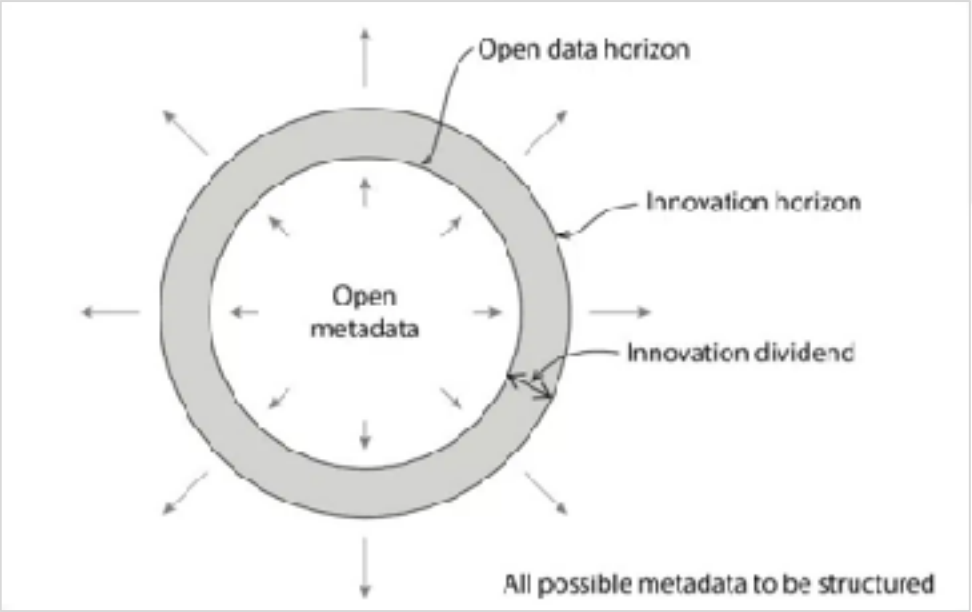


Latin America Research Funding Patterns- A tutorial using the Dimensions on Google BigQuery
https://colab.research.google.com/drive/1iqPNkJjCTbjPMZvk2C9qt8rZYSQD8IZU

Dimensions

Openness: what's open, what isn't

- Dimensions is a **demonstration** of what can be done when data are made freely and openly available.
    - CrossRef, DataCite, PubMed, ORCID, arXiv etc..
- We are researchers, we build for researchers
    - Free access to Dimensions for scientometrics projects
    - 2017: GRID institution data CC0 (=> ROR)
    - Grant data freely available to funders with < $1m in funding per year
    - ANZSRC and SDG classification freely available in Dimensions;
- Key **principle**: Striking the right balance between enabling entrepreneurial spirit and activity and supporting open science and open data e.g. Barcelona Declaration



Daniel Hook, 2024, Barcelona: A beautiful horizon., https://www.digital-science.com/tldr/article/barcelona-a-beautiful-horizon/

DIGITAL science

Introducing SRAD

Our Scientometric Researcher Access to Data (SRAD) program makes data available at no charge for non-commercial scientometric research projects.

Inclusive datasets like Dimensions are essential for understanding the global research landscape. We are proud to make these data available for the study of science, to support innovative research that investigates patterns and trends in research and its influence on science, policy, innovation, and communication.



Open questions for the panel

- For granting councils:
  - Which funders and agencies should be in Dimensions that aren't? What would it take e.g. a data submission problem, a standards problem, a trust problem?
- For access:
  - What would meaningful access to funding flow data look like for Southern granting councils that can't afford subscriptions?
- For governance:
  - The Barcelona Declaration is building consensus around open research information. What role should commercial infrastructure providers play — and on what terms?



Thanks

References

Dimensions Data:

- White paper: A Guide to the Dimensions Data Approach (2023)
- Dimensions Knowledge Base

Dimensions on BigQuery:

- Online documentation: https://docs.dimensions.ai/bigquery/
- GBQ LAB: https://bigquery-lab.dimensions.ai/index.html

Dimensions API (DSL):

- Online documentation: https://docs.dimensions.ai/dsl
- API LAB: https://api-lab.dimensions.ai/