Report

Infrastructure and Awareness Landscape Analysis in Latin America

Arturo Garduño-Magaña iD

DataCite is a global community that shares a common interest: to ensure that research outputs and resources are openly available and connected so that their reuse can advance knowledge across and between disciplines, now and in the future. As a community, we make research more effective with metadata that connects research outputs and resources–from samples and images to data and preprints. We enable the creation and management of persistent identifiers (PIDs), integrate services to improve research workflows, and facilitate the discovery and reuse of research outputs and resources.

DataCite is registered as a non-profit organization in Hannover, Germany, with a team distributed across the globe. For any questions or comments about the report, please contact us at: support@datacite.org.

Cover image by kjpargeter on Freepik.

DataCite Regional Landscape Analysis: Latin America

1. Introduction

This landscape analysis examines the current state of infrastructure and awareness regarding the adoption of Persistent Identifiers (PIDs) in Latin America. PIDs are crucial for ensuring the long-term accessibility, discoverability, and citability of research outputs and resources. Understanding the existing infrastructure and awareness levels is essential for assessing the region's readiness and identifying potential areas of improvement.

PIDs allow for the unique and persistent identification of research results and resources, whether they are digital objects such as datasets, manuscripts, or software, or even physical objects like instruments or samples, among others. This simplifies the search, referencing, and citation in the field of science. Effectively identifying and sharing data increases the visibility of research and promotes its reuse. Thus, accelerating the advancement of science and strengthening the relationship between research and society.

Latin America, with its diverse and dynamic research community, is at the forefront of embracing open science principles. This landscape analysis aims to provide insights into the existing infrastructure and awareness levels in the region, recognizing that a nuanced understanding of the current state of affairs is essential for assessing the region's readiness and identifying challenges and opportunities .

As the research landscape in Latin America evolves, it is crucial to recognize that the adoption of PIDs is not merely a technical matter but a transformational step towards ensuring that research outputs are preserved, accessible, and influential. By examining the current status, we aim to foster a collaborative and informed discussion on how Latin America can harness the full potential of PIDs and open science to engage more robustly with the global academic community.

This analysis seeks to highlight the progress made in the region and to identify areas where investments, support, and collaboration can further accelerate the adoption of PIDs and open science. Ultimately, it aims to be a first step on how to better understand the regional needs and challenges, to help the organization to support

local communities to build a more open and equitable research infrastructure and best practices.

2. Scope and Methodology

This report explores the research landscape of the Latin American region to gain deeper insights into the underlying research infrastructure. The report focuses on two major aspects: A) Infrastructure landscape analysis and B) Awareness landscape analysis. For each aspect, we detail the scope, methodology, and data sources. This study aims to highlight the nuances of infrastructure and awareness levels, highlighting the prospects and obstacles within the Latin American context.

2.1 Infrastructure Landscape Analysis

Methodology and Data Sources

The following standardized and internal data sources and systems were used:

- Registry of Open Access Repositories (ROAR)
- Directory of Open Access Repositories (OpenDOAR)
- Confederation of Open Access Repositories
- Registry of Research Data Repositories (Re3data.org)
- Public Knowledge Project Usage Data
- Dataverse Installations Data
- DataCite Fabrica

The data was gathered manually by cross-referencing each country with multiple sources. In certain instances, the raw data sheets were refined to meet the report's specific requirements. Subsequently, all of this data was processed, and charts were generated to provide insights into the infrastructure landscape in Latin America.

Coverage

Within the infrastructure landscape analysis, we cover the following information and data for 23 countries in the Latin American region:

- Number of repositories per country

- Service Providers (SP) per country
- Latin American Repository Aggregators and Networks
- Number of Dataverse installations per country
- Number of DSpace installations per country
- Number of Public Knowledge Project (PKP) installations for the following products: Open Journal Systems (OJS), Open Monograph Press (OMP), Open Preprint Server (OPS)
- Current PID Systems in Latin American repositories
- Current status adoption and membership of DataCite in Latin America

2.2 Awareness Landscape Analysis

Methodology and Data Sources

Within the awareness landscape analysis, we cover the following points:

- Existing institutional policies among open access repositories across Latin American institutional policies
- Landscape of National Open Science policies in Latin America
- Current DataCite memberships and adoptions across Latin America
- DataCite Community in the region: Expert groups, and board members
- Individual interviews that we did with the community

Data Collection

As for the data sources, the following standardized and internal sources were used:

- **Registry of Open Access Repositories (ROARMAP)**: This source provides valuable insights into existing institutional policies among open access repositories.
- **DataCite Internal Data**: Offers information on current DataCite memberships and adoptions in Latin America.
- **DataCite Fabrica**: Utilizing internal DataCite systems such as Fabrica helps in understanding the DataCite community in the region

Coverage

In this landscape analysis, we examine the following aspects within the Latin American context:

- The institutional policies of open access repositories throughout Latin America.
- Current memberships and the adoption status of DataCite within the region.
- The DataCite Community in Latin America, encompassing board members and expert groups.

Countries

Both for the Infrastructure Landscape Analysis and the Awareness Landscape Analysis aspects, only countries with reported data in the aforementioned data sources were included. Of the 33 Latin American countries according to CEPAL[1], only 23 have data available in the sources consulted. These countries are:

Country		
Argentina	Dominican Republic	Nicaragua
Bolivia	Ecuador	Panama
Brazil	El Salvador	Paraguay
Chile	French Guyana	Peru
Colombia	Guatemala	Puerto Rico
Costa Rica	Honduras	Uruguay
Cuba	Jamaica	Venezuela
Dominica	Mexico	

Table 1. List of countries in the Latin American Region included in the analysis

3. Infrastructure Landscape Analysis in Latin America

Persistent Identifiers (PIDs) are essential tools in the digital research ecosystem, providing unique and long-lasting references for various types of research outputs

---

[1] See https://www.cepal.org/en/about/member-states

and resources. Among the different types of PIDs, Digital Object Identifiers (DOIs) are the most widely recognized and implemented globally. DOIs ensure that scholarly articles, conference papers, datasets, and other research materials can be easily discovered, cited, and accessed.

3.1 Current state of existing PID infrastructure in Latin America

When it comes to the infrastructure for Persistent Identifiers (PIDs), Latin American countries have been making notable strides, including in the adoption of Digital Object Identifiers (DOIs). DOIs are internationally recognized and extensively implemented. They serve as a fundamental tool for uniquely identifying various types of research outputs and resources, including scholarly articles, conference papers, and datasets.

In Latin America, organizations like DataCite play a pivotal role by offering DOI registration services. This advancement signifies a commitment to ensuring that research outputs from the region are easily discoverable, citable, and globally accessible through the use of standardized PIDs, specifically DOIs. These efforts strengthen the region's position in the global research landscape, promoting collaboration and knowledge exchange.

3.2 Current state of DataCite consortia in Latin America

DataCite plays a crucial role in supporting the scholarly community by providing DOI registration services and fostering the adoption of PIDs. One of the key strategies for extending this support globally is through the formation of consortia. These consortia bring together research institutions, libraries, and other stakeholders to collaborate on best practices for managing and sharing research data.

DataCite's mission is to make research outputs easily discoverable, accessible, and citable. By establishing consortia in Latin America, research institutions, libraries, and other stakeholders in the region can ensure that the valuable research produced there is accessible not only locally but also globally. This facilitates international collaboration and access to research outputs, benefiting both Latin American researchers and their global counterparts.

The already existing DataCite consortia in the region play a pivotal role in providing guidance, training, and best practices for managing and sharing research data, which, in turn, can lead to an increase in data quality and compliance with international standards.

DataCite has a total of 54 consortium organizations distributed across 6 consortia in 5 different countries (see Fig. 1). The Colombian consortium, Consortia SAS, has the highest number of consortium organizations. The Mexican consortium, eScire, leads in terms of the highest number of repositories, and the Brazilian consortium, CNPq, leads in the number of registered DOIs, according to DataCite's internal data.
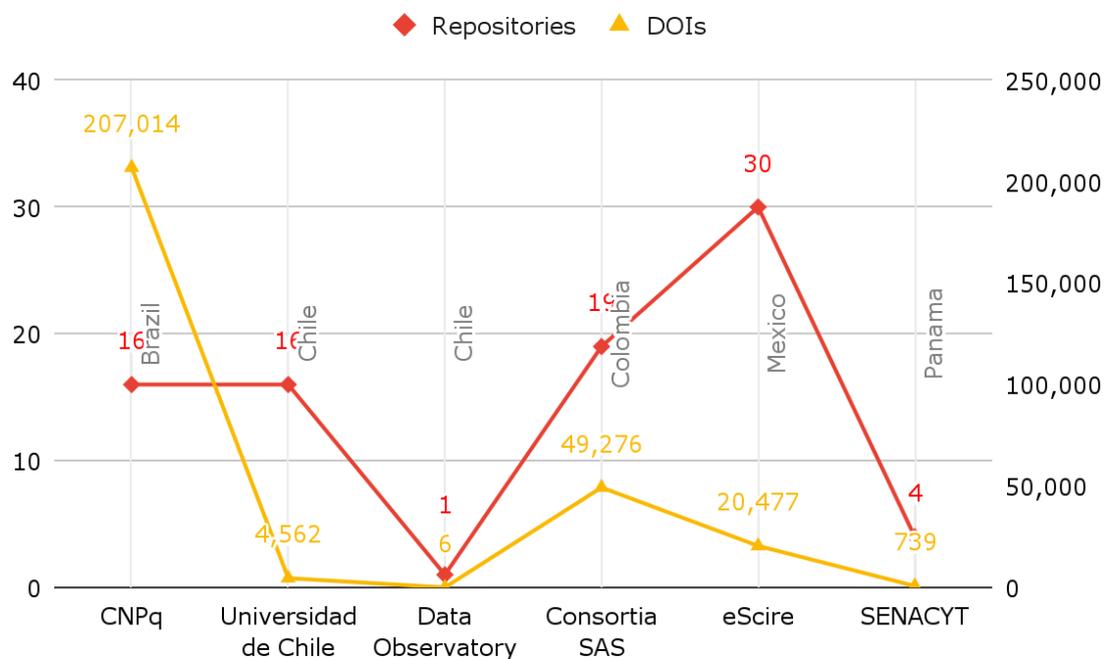

Fig. 1 Number of repositories and registered DOIs by consortium, by October 2024

Figure 1 displays the number of repositories and registered DOIs by consortium in Latin America, as per DataCite Fabrica. As previously mentioned, Mexico takes the lead in the number of repositories, with Colombia and Brazil following closely behind. It's worth noting that the Chilean consortium, Data Observatory, is the most recent addition to DataCite and has recently established its own repository, along with registering their first DOI.
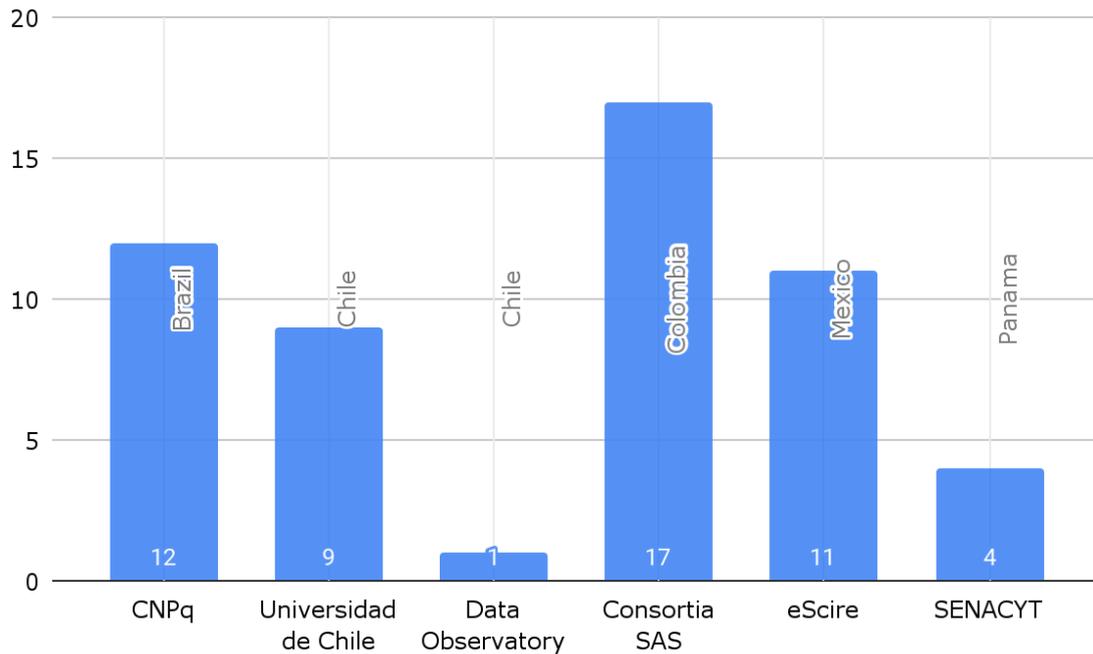
Fig. 2 Number of members per consortium

It is important to note that, in addition to the existence of established consortia in the region, there are two organizations in Argentina and Ecuador who have joined the Dataverse Consortium due to their strong affiliation with the Dataverse community. Additionally, there is one direct member in Brazil, as presented below.

Country	Organization	Consortium	Repositories	DOIs
Brazil	INSPER	Direct Member	1	86
Argentina	Universidad Nacional del Rosario	The Global Dataverse Community Consortium (GDCC)	1	263
Ecuador	CEDIA	The Global Dataverse Community Consortium (GDCC)	1	12
Guatemala	Universidad del Valle de Guatemala	The Global Dataverse Community Consortium (GDCC)	1	
Uruguay	Agencia Nacional de Investigación e Innovación (ANII)	The Global Dataverse Community Consortium (GDCC)	1	

Tab. 2 DataCite members not affiliated to Latin American consortia
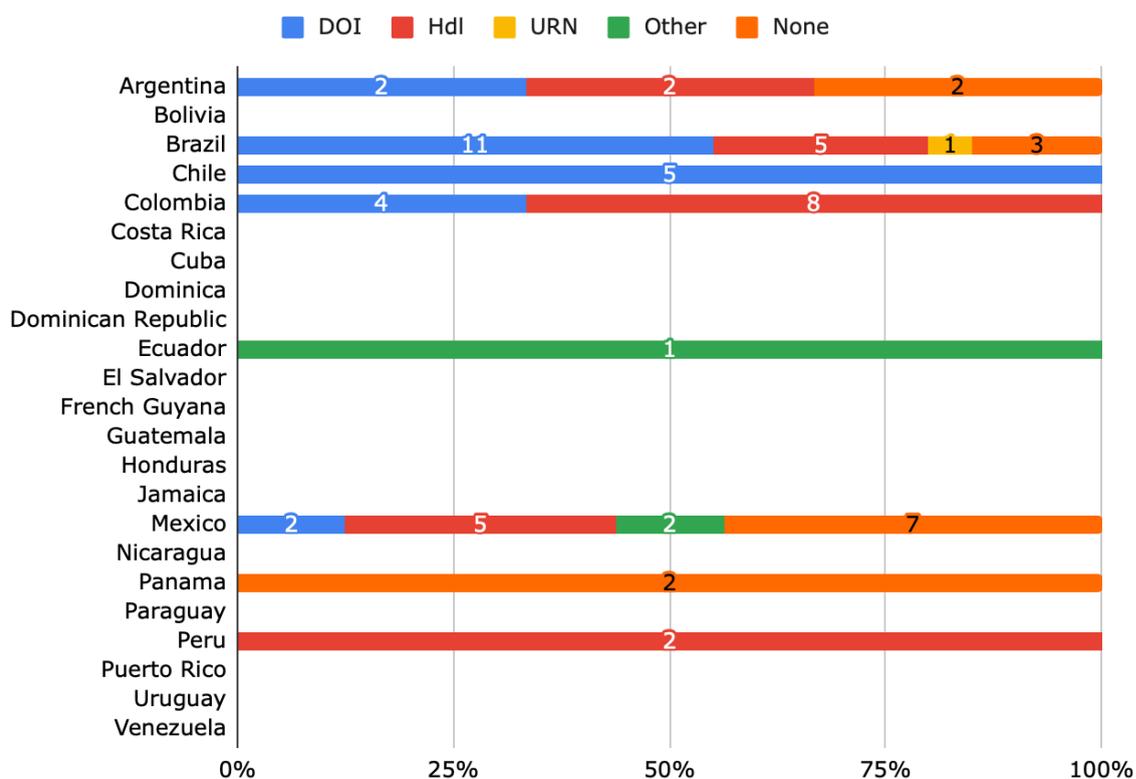
3.3 Adoption of PIDs in Latin America

Fig. 3 Comparison of PID systems adoption in repositories. Source: Re3data

Data Repositories play an important role in DataCite's mission to cite, provide access and reuse research data. Figure 3 illustrates a comparison of the choice of PID systems and their adoption in data repositories by country within the region, as documented by Re3data. The distribution shows heterogeneity, with clear preferences in certain countries. For example, Chile predominantly uses DOIs, while Peru prefers Handle (Hdl). It's also worth noting the cases of Argentina and Mexico, where data repositories are registered using a combination of DOIs, Hdl, or other identifiers.
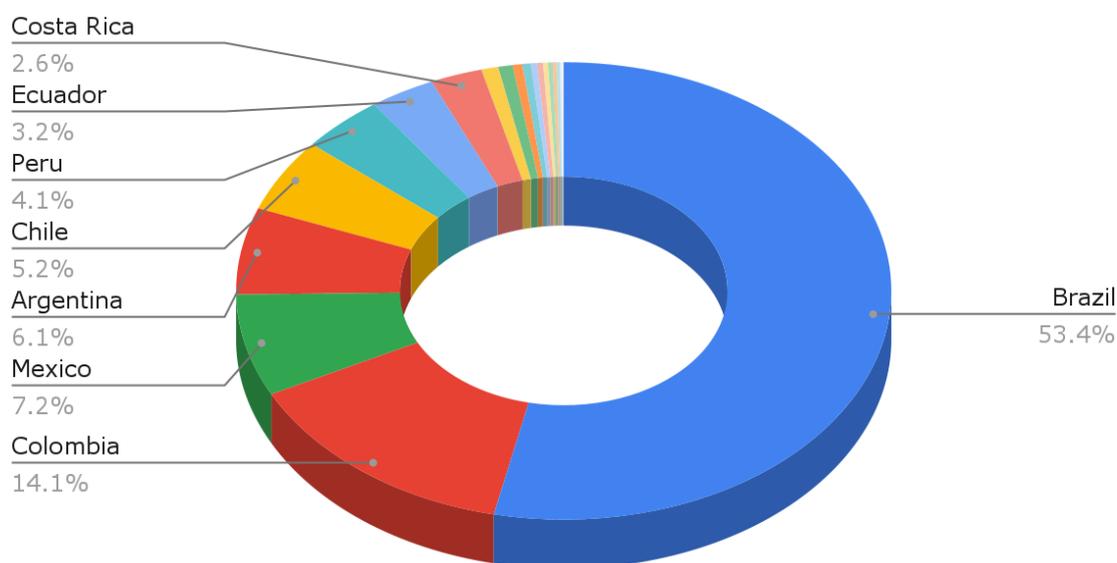
Fig. 4 Distribution of journals using DOI as their PID system. Source: Directory of Open Access Journals

Open Access Journals also have a significant history of adopting and implementing PIDs. Regarding DOIs, the Directory of Open Access Journals (DOAJ) offers valuable insights into the distribution of DOI adoption as the preferred PID by country in the region (Fig. 4). It's evident that Brazil leads with the highest adoption rate, followed by Colombia, Mexico, and Argentina once again.

3.4 Current state of infrastructure and repositories networks in the region

Open access to scientific research and the development of community-managed open infrastructures and systems have been a consistent theme in the region due to the high costs associated with accessing paid resources and tools. Repositories have served as excellent tools for not only storing but also preserving, disseminating, and increasing the visibility of Latin American research.

Several countries have implemented institutional and national policies to formalize and strengthen the workflows, management, and standardization of the existing

institutional and thematic repositories. However, it is challenging to identify all the repositories and information management systems, as most of the international directories and registers are self-declared.

The following numbers and comparisons are presented based on registries, directories, and national aggregators identified in Latin America.
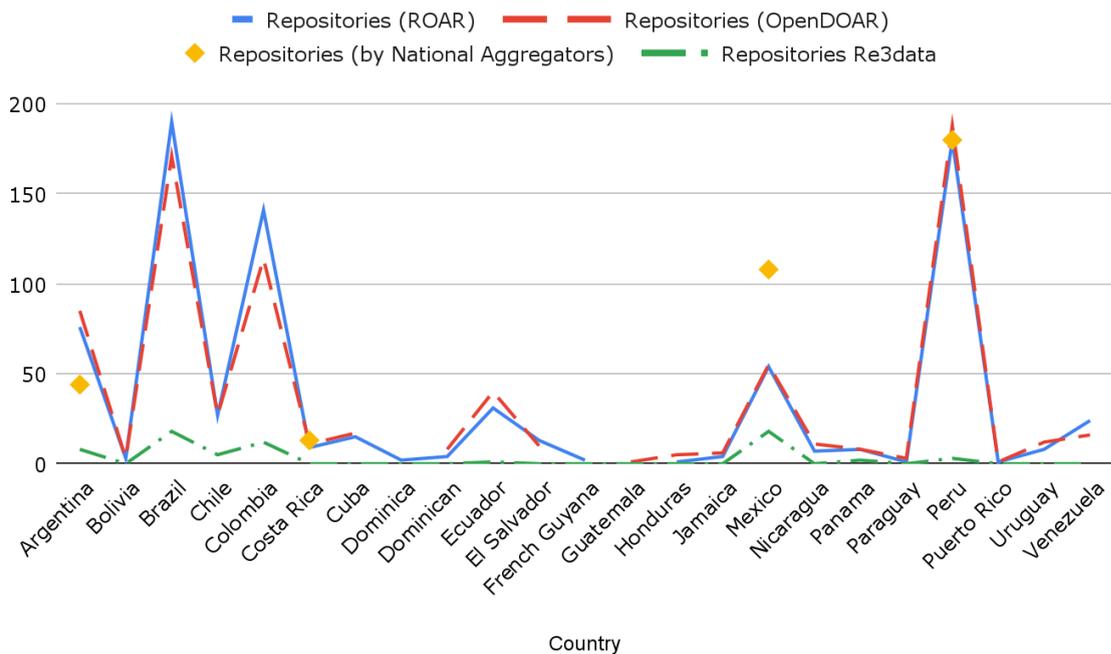

Fig. 5 Number of Repositories comparison by country. Sources: Registry of Open Access Repositories, Directory of Open Access Repositories, Re3data, and National Aggregators

Figure 5 illustrates a comparison of the number of repositories registered in ROAR, DOAR, Re3data, and the identified national aggregators. It's worth noting the case of Peru, which not only boasts one of the highest numbers of self-reported repositories but also demonstrates a particularly homogeneous comparison among the registries. Brazil and Colombia also stand out, sharing the lead with Peru as the countries with the highest number of self-reported repositories.
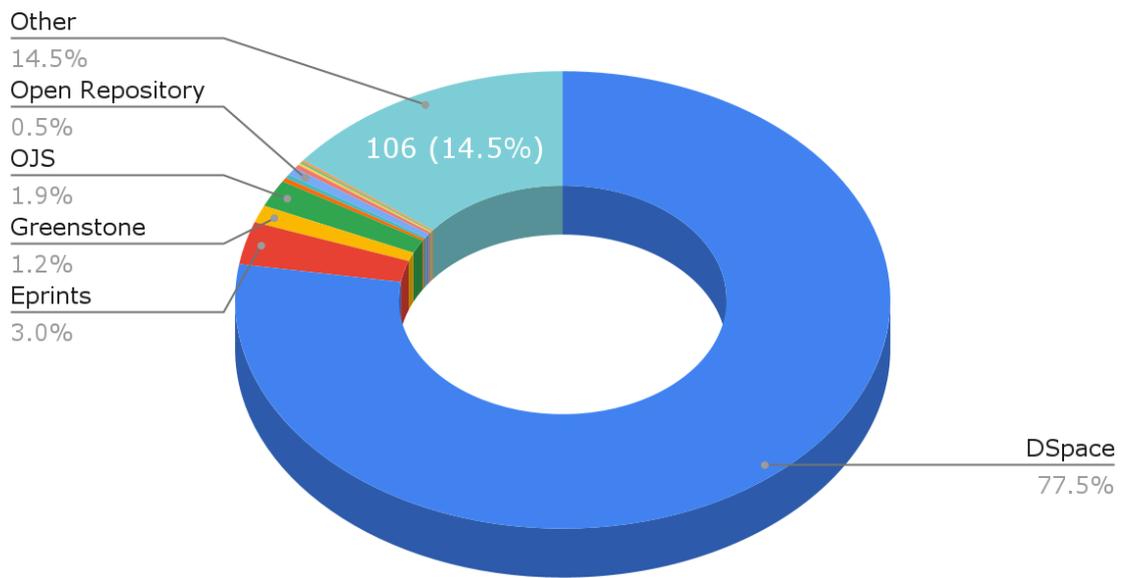
Other
14.5%

Open Repository
0.5%

OJS
1.9%

Greenstone
1.2%

Eprints
3.0%

106 (14.5%)

DSpace
77.5%

Fig. 6 Repositories Service Providers, Registry of Open Access Repositories

Other
8.8%

DSpace CRIS
1.2%

Cybertesis
2.7%

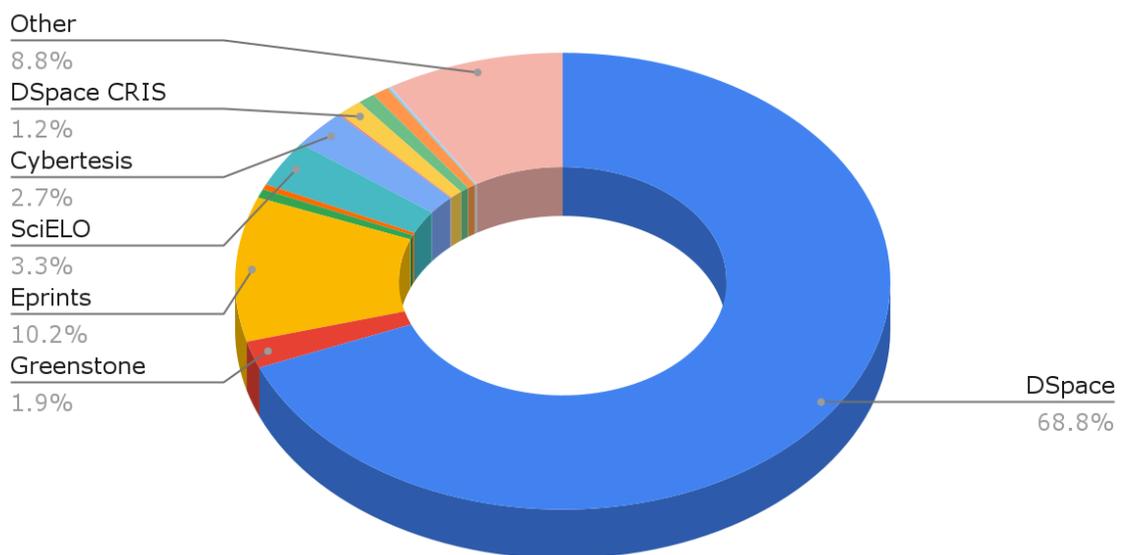SciELO
3.3%

Eprints
10.2%

Greenstone
1.9%

DSpace
68.8%

Fig. 7 Repositories Service Providers, Directory of Open Access Repositories
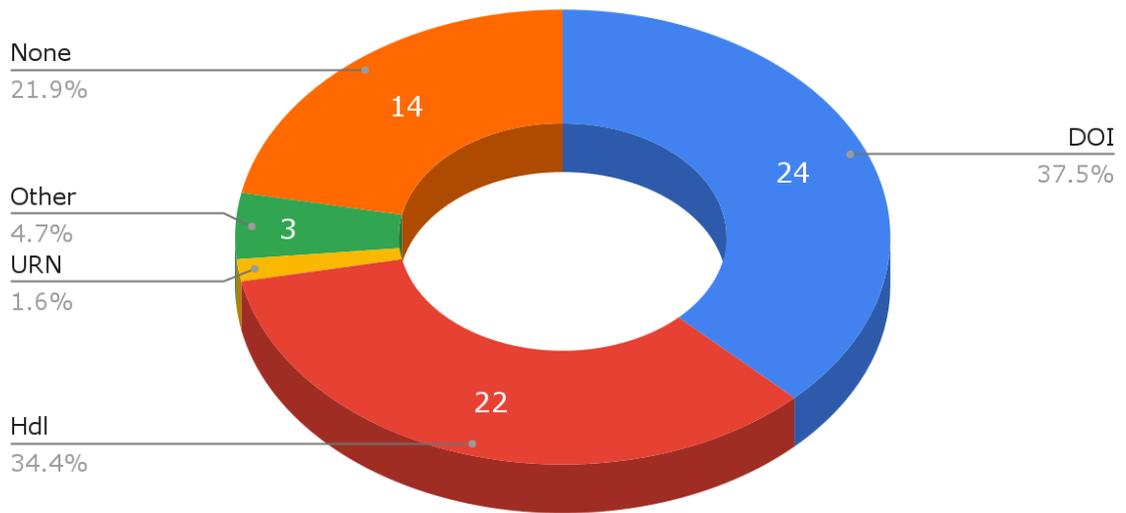
Fig. 8 Repositories Service Providers, Re3data

Figures 6, 7, and 8 illustrate the diversity in the adoption of repository service providers. DSpace stands out as the primary choice for repositories registered in ROAR and DOAR. In contrast, for data repositories registered in Re3data, DSpace is the second most common choice, surpassed only by Dataverse. Dataverse is a widely popular service provider for data, software, and other repositories, and its usage is steadily increasing.
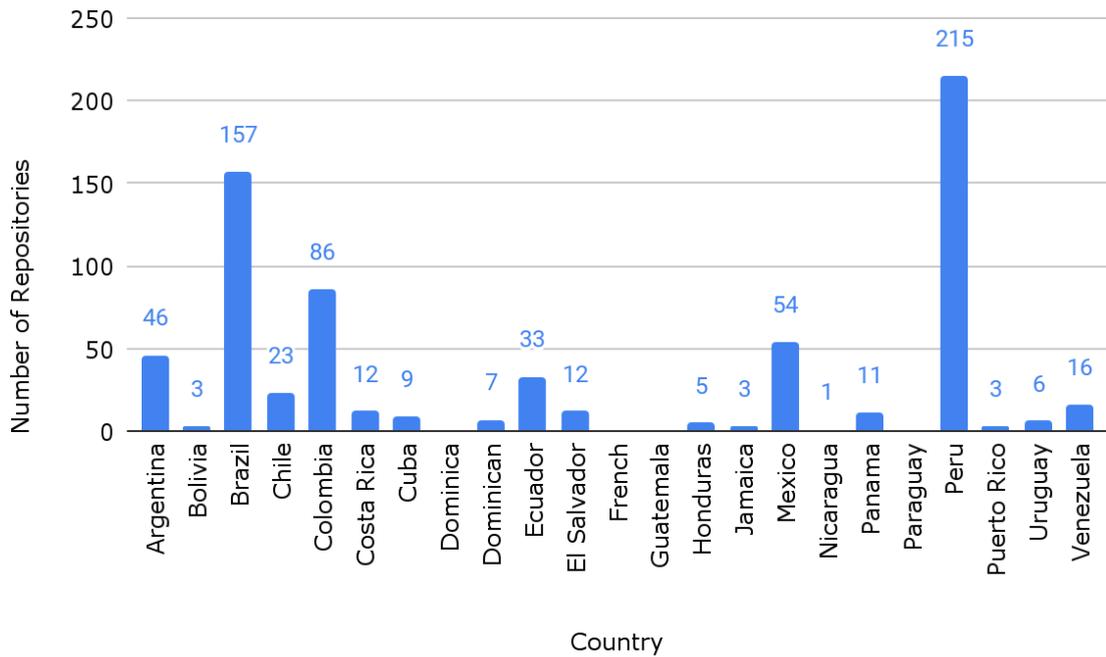
Fig. 9. DSpace installations per country in Latin America, Lyrasis

As mentioned earlier, DSpace is the primary choice for institutions' repositories in Latin America. Figure 9 illustrates the number of self-reported DSpace installations, according to Lyrasis, the organization behind DSpace. As you can see, Peru once again leads the comparison in the region. By comparing it to Figure 5, it can be deduced that Peru has embraced the culture of self-reporting in international registries and directories most effectively. Brazil and Colombia consistently follow Peru in the top positions.
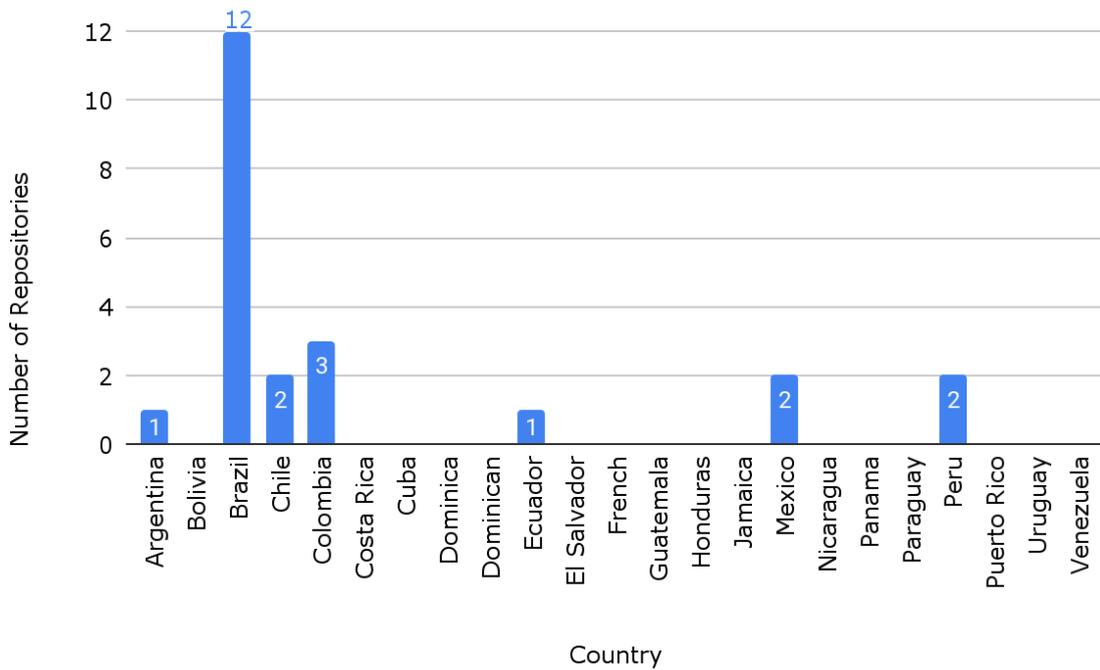
Fig. 10. Dataverse installations per country in Latin America, Dataverse

Although the adoption of Dataverse in the region has been growing in recent years, according to the organization's own website, as of September 2023, there are 23 self-reported Dataverse installations as a service provider for repositories in the region. Brazil has the highest number of registered installations, followed by Colombia, Chile, Mexico, and Peru, with Argentina and Ecuador also having recorded installations, albeit to a lesser extent.

3.5 The relationship between Public Knowledge Project and DataCite

The Public Knowledge Project (PKP) is a non-profit organization that develops open-source software used by academic journals and other publications for the management and online publication of academic content.

The relationship between PKP and DataCite is manifested in the ability of PKP's publishing systems, such as Open Journals Systems (OJS), to assign DOIs to academic publications and other digital objects they manage. This collaboration enables academic journal editors and other providers of academic content to utilize

DataCite's services to ensure that their publications are easily citable and traceable within the academic and research community. PKP is also a certified Service Provider for DataCite.
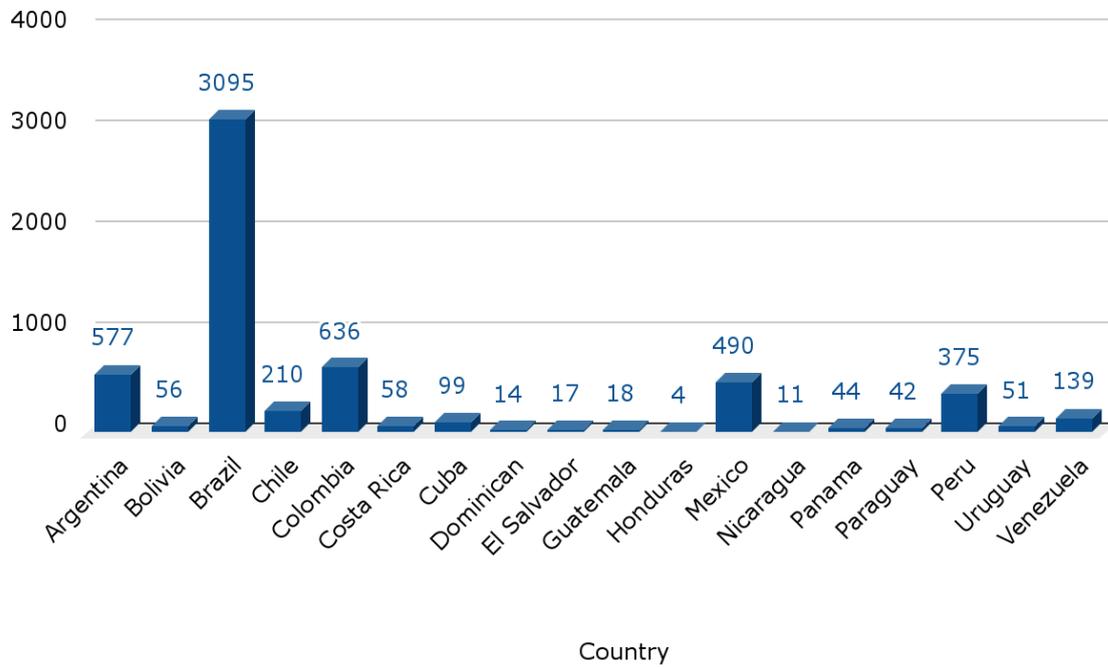

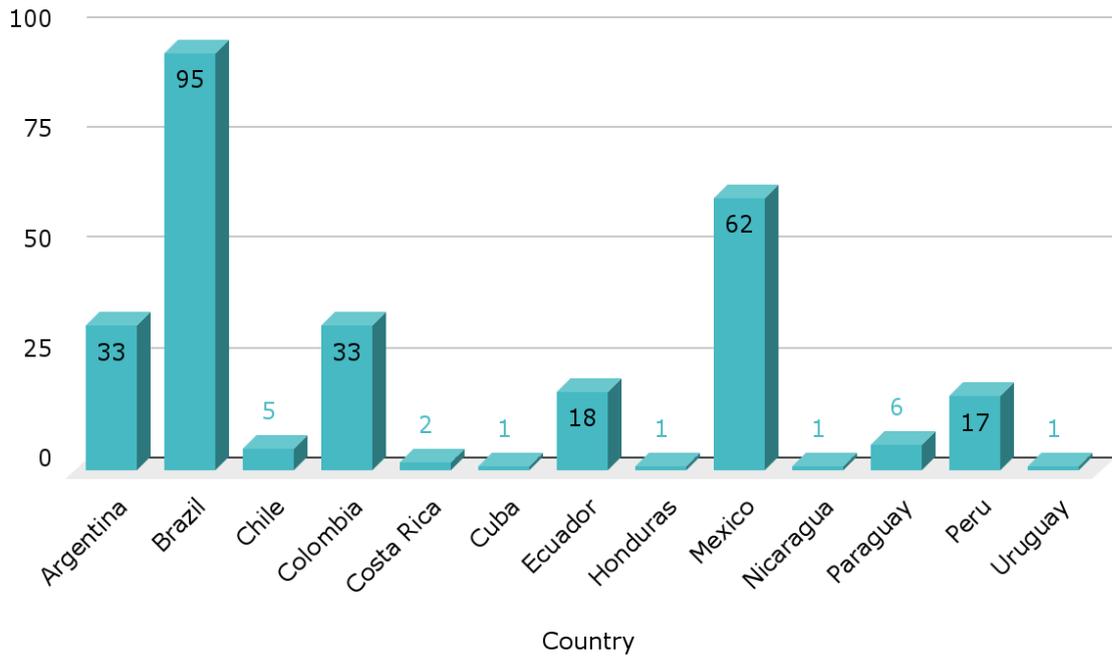Fig. 11 Open Journal System Installations per country, PKP

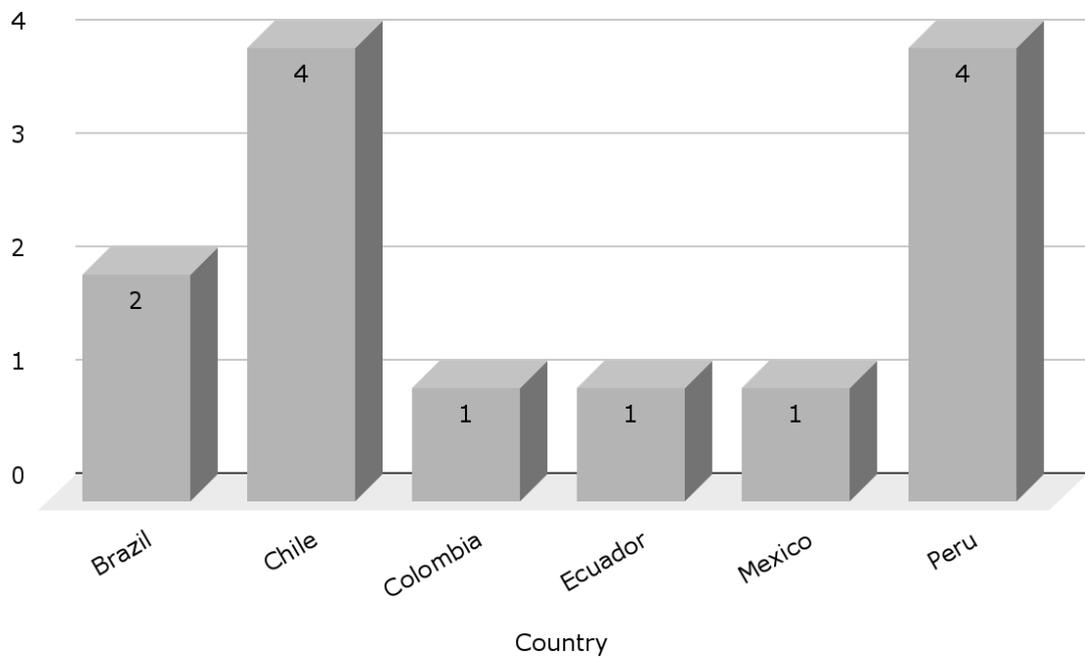Fig. 12 Open Monograph Press Installations per country, PKP

Fig. 13 Open Preprint System installations per country, PKP

As shown in Figures 11, 12, and 13, the adoption of OJS in the region holds significant value. This is especially true because universities and research centers, often public institutions, support open-access journals and access to research outputs in Latin America. This has compelled them to discover, adopt, and leverage one of the most versatile tools for managing academic publications. The trend continues, with Brazil once again leading as the country with the highest number of OJS installations, followed by Colombia in second place, and Argentina in third.

This is similar for OMP, although with a smaller number of self-reported installations. Brazil and Mexico lead the chart, followed by Argentina and Colombia.

In the case of OPS (Open Preprint Systems), which is a more recent product from PKP, there isn't enough data to demonstrate a clear trend. However, we can observe that Chile and Peru have been the most enthusiastic in reporting installations of the tool in the region.

4. Policies and Strategies on Open Science in Latin America

4.1 National and Institutional Open Science Policies in Latin America

Across Latin America, there is a growing momentum in the adoption of Open Science policies and the integration of best practices throughout the research cycle. Several countries in the region have taken significant steps in this direction. For instance, countries like Peru (since 2013), Argentina (since 2013), Mexico (since 2014), and Colombia (since 2022) have implemented national laws to support open access repositories and have introduced comprehensive national plans for Open Science.

Countries like Brazil, Costa Rica, Cuba, Chile, and other nations in the region have been actively working on the creation of various Open Science tools, including

manifestos, official declarations, strategic action plans, and national platforms for open access publications[2].

Fig. 14 Countries with Open Science Policies, UNESCO

The regional map (Fig. 14) displays the study conducted by UNESCO, which records the national policies and strategies related to Open Science in Latin America. It's worth noting that a significant portion of the region has made progress in this regard. However, there is also a noticeable absence of monitoring and follow-up mechanisms to assess and update these strategies and policies.

---

[2] Heredia, Ana. "A Tradition of Open, Academy-Owned, and Non-Profit Research Infrastructure in Latin America." Information Services & Use, edited by Bonnie Lawlor, vol. 42, no. 3–4, Dec. 2022, pp. 447–52, https://doi.org/10.3233/ISU-220177.

datacite.org

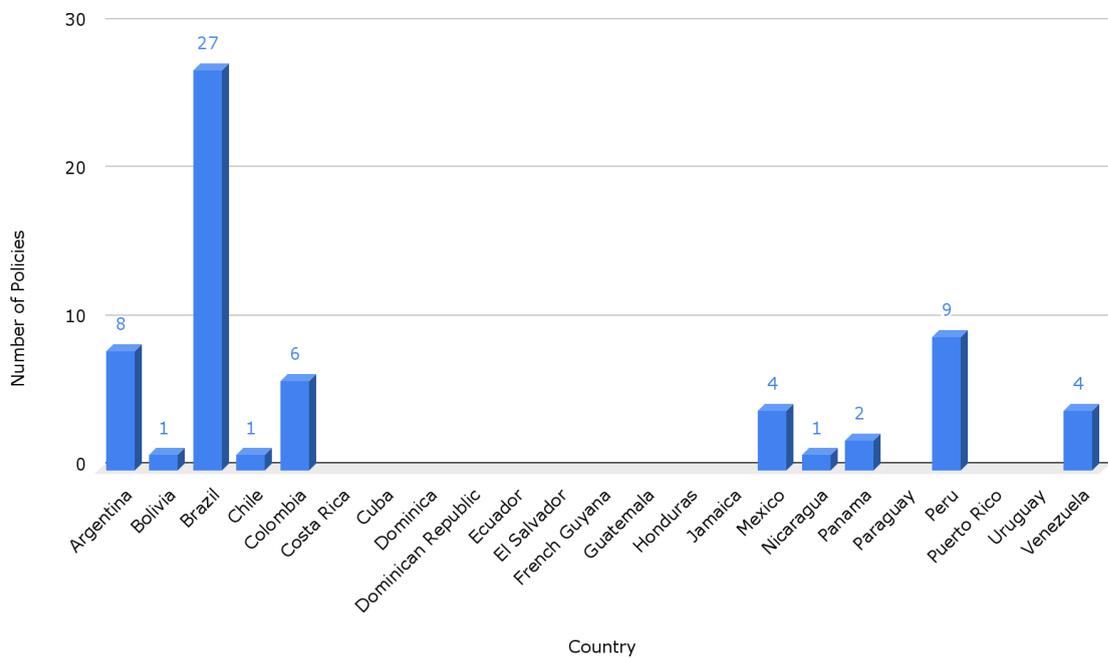## 4.2 Mandatory Archiving Policies per Country in Latin America

Fig. 15. Number of Mandatory Archiving Policies per Country, ROARMAP

Figure 15 presents data sourced from the Registry of Open Access Repositories Mandates and Policies (ROARMAP), showing the count of mandatory archiving policies per country in Latin America. Brazil leads with the highest number of registered mandatory archiving policies, followed by Peru and Argentina. While it's important to note that this registry relies on self-reporting, it's evident that several countries in the region have yet to provide information on this matter. This underscores the need for continued efforts to enhance transparency and accountability in the region, leaving room for improvement in reporting and compliance.

Policymaker
1.6%

Funder and
3.1%

Sub-unit of
12.5%
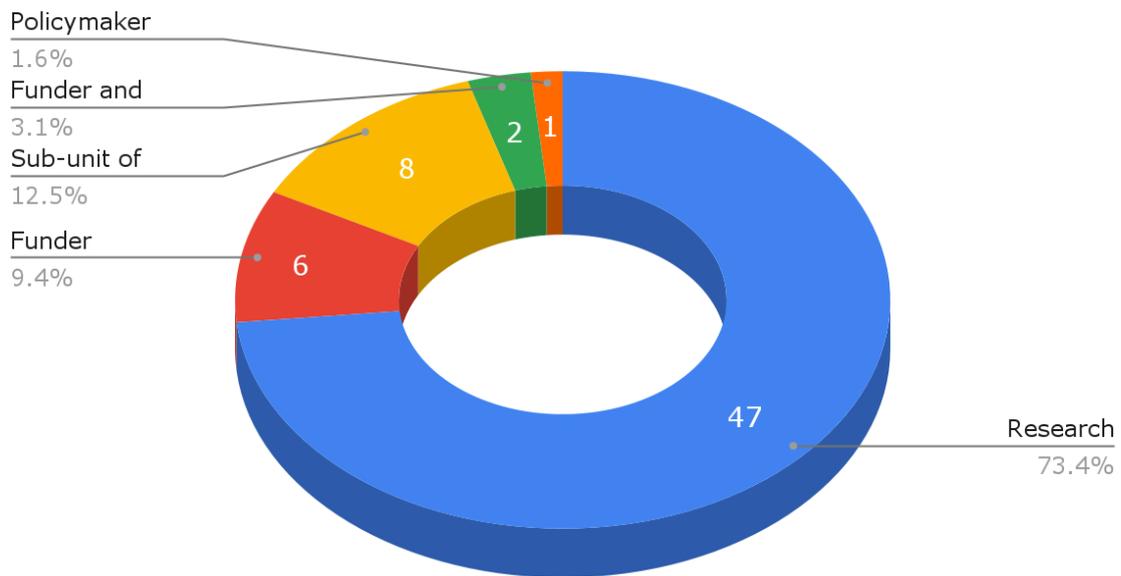
Funder
9.4%

Research
73.4%

Fig. 16. Distribution of Types of Organizations with Mandatory Archiving Policies, ROARMAP

Figure 16 illustrates the distribution of the types of organizations that have implemented the mandatory archiving policies reported by ROARMAP. Research organizations have reported the highest number of policies, followed by research sub-units. It is important to note that Funders and Policymakers have reported a lower number, which may not necessarily imply their absence but rather a potential lack of monitoring regarding the implementation of these policies.

The adoption of Open Science policies in Latin America is a topic of growing interest and discussion. As mentioned earlier, research conducted in the region indicates that there is a notable awareness of Open Science practices. However, the actual implementation of these policies remains a challenge.

Only a limited number of Latin American journals have successfully incorporated Open Science practices into their operations. Although there are promising indications of Open Science initiatives emerging in various countries, especially with a strong emphasis on improving open access infrastructures, there is still more progress to be made. It is evident that there is a solid foundation of awareness and some headway in adopting Open Science practices and policies, but further efforts

are needed to promote increased implementation and cultivate a culture of openness in research and academia within the region.

5. Discussion

This analysis highlights significant progress and challenges in the Latin American research landscape, particularly regarding the adoption of PIDs and Open Science policies. The region demonstrates a strong commitment to open access and research visibility, driven by regional entities and consortia. Organizations like DataCite and local consortia have catalyzed adoption and ensured that Latin American research aligns with global standards. However, challenges persist, especially regarding the fragmentation of PID systems and the need for better infrastructure support across all countries.

The diversity in PID adoption across countries underscores the need for a more unified approach. For example, while some countries favor DOIs, others utilize a mixture of systems, which can create interoperability challenges. Latin America's unique position, balancing local research priorities with global integration, demands tailored support that fosters both regional consistency and international collaboration. Additionally, as seen in the variation of Open Science policies, there is an evident gap in systematic monitoring, which could hinder sustained progress.

The infrastructure for repositories also illustrates Latin America's commitment to open access, yet there remains room for improvement in terms of self-reporting accuracy and standardization. Countries like Peru and Brazil lead in repository numbers, showcasing their engagement with open science infrastructure, but smaller countries face significant barriers. Addressing these disparities requires targeted policies that consider the varying capacities across the region.

6. Conclusion

Latin America has laid a strong foundation for open science through the adoption of PIDs, the formation of consortia, and the establishment of Open Science policies.

The region's progress in integrating PIDs and developing national policies points to a promising trajectory. Nonetheless, this analysis reveals critical areas for development: fostering a more unified adoption of PID systems, enhancing infrastructure support, and implementing robust monitoring mechanisms to ensure compliance with open science mandates.

A comprehensive approach that incorporates these elements will allow Latin America to not only address existing gaps but also position itself as a leader in global open science practices. The continued collaboration among regional institutions, support from global entities, and a commitment to policy refinement will be key to sustaining the momentum and achieving long-term impact in the research landscape of Latin America.

7. References

Heredia, Ana. "A Tradition of Open, Academy-Owned, and Non-Profit Research Infrastructure in Latin America." Information Services & Use, edited by Bonnie Lawlor, vol. 42, no. 3–4, Dec. 2022, pp. 447–52, https://doi.org/10.3233/ISU-220177.

Mejias, G., & Cousijn, H. (2023). DataCite launches Global Access Program with support from CZI. https://doi.org/10.5438/3GS5-3J26

United Nations Economic Commission for Latin America and the Caribbean (CEPAL). (n.d.). Member States. Retrieved from https://www.cepal.org/en/about/member-states

DataCite

CONNECTING RESEARCH,
ADVANCING KNOWLEDGE

www.datacite.org