

World Data System

Research Analysis: WDS Members and Canadian CTS Cohort Needs Assessment

14 March 2023

Caroline Lee
Research Associate for International Technology Office
ito-ra3@oceannetworks.ca

Sarah Gonzalez
Program Manager for International Program Office
sgonzal4@utk.edu

Karen Payne
Director for International Technology Office
ito-director@oceannetworks.ca



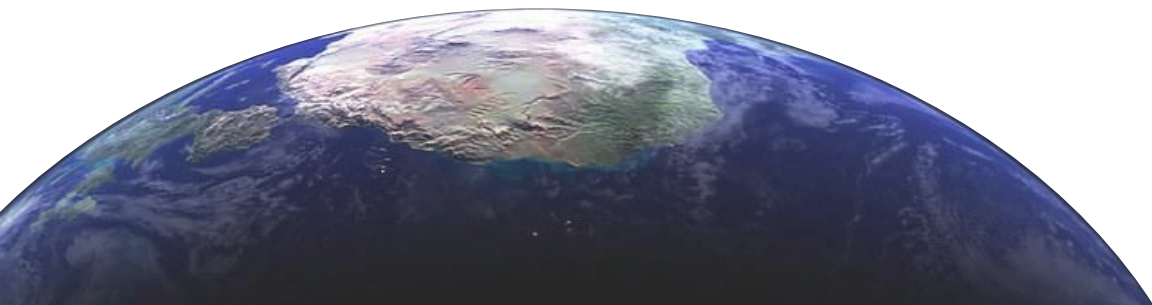
Overview



From July 2022 to December 2022, the World Data System (WDS) International Technology (ITO) and International Program (IPO) Offices **conducted a review of strategic plans and technical roadmaps to identify needs and challenges** for:

- WDS Members ([Regular, Network, Associate and Partner](#))
- The Canadian repositories participating in the [Digital Research Alliance of Canada's CoreTrustSeal Certification Support and Funding Pilot](#)

Additionally, incorporated into this assessment was criteria from the [Global Biodata Coalition \(GBC\)](#), which is an initiative created to support funders by identifying and prioritizing fundamental data resources (Global Biodata Coalition, 2022).

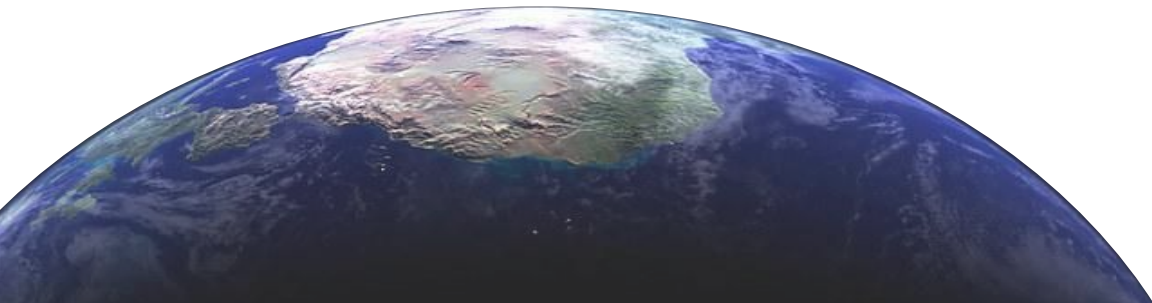


Overview

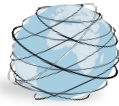


The **goal** of this assessment was, first and foremost, to identify the needs and challenges faced by WDS and Canadian CTS member repositories. Additional goals include:

- Support repositories with creating guiding architecture plans for integrating with other global research services
- Help prioritize funding to projects that support common infrastructure needs
- Help members who seek to be represented in collaborative research environments of choice
- Identify WDS datasets that support polar research
- Identify which members have satisfied [CoreTrustSeal](#) requirements



Overview



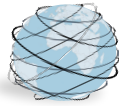
A **reviewer instrument** was designed that enabled an organized and thorough review of all documentation, incorporating criteria designed to identify focus areas and priorities for data repositories and the GBC criteria.

- We augmented our initial criteria with the GBC criteria to create the final reviewer instrument that was implemented using google forms.

Reviews of the collection of strategic plans commenced in October 2022. At this time, each document has been reviewed once.

- The reviewer instrument was tested with five strategic plans collected from three WDS regular members, and two Canadian CTS Cohort members.

Overview



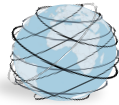
A **reviewer instrument** was designed that enabled an organized and thorough review of all documentation, incorporating criteria designed to identify focus areas and priorities for data repositories and the GBC criteria.

- We augmented our initial criteria with the GBC criteria to create the final reviewer instrument that was implemented using google forms.

Reviews of the collection of strategic plans commenced in October 2022. At this time, each document has been reviewed once.

- The reviewer instrument was tested with five strategic plans collected from three WDS regular members, and two Canadian CTS Cohort members.

Overview



A review
review
areas

13. Mission: Does the plan include a mission statement?

☐ Yes

☐ No

☐ Maybe

Quantitative

Review
time,

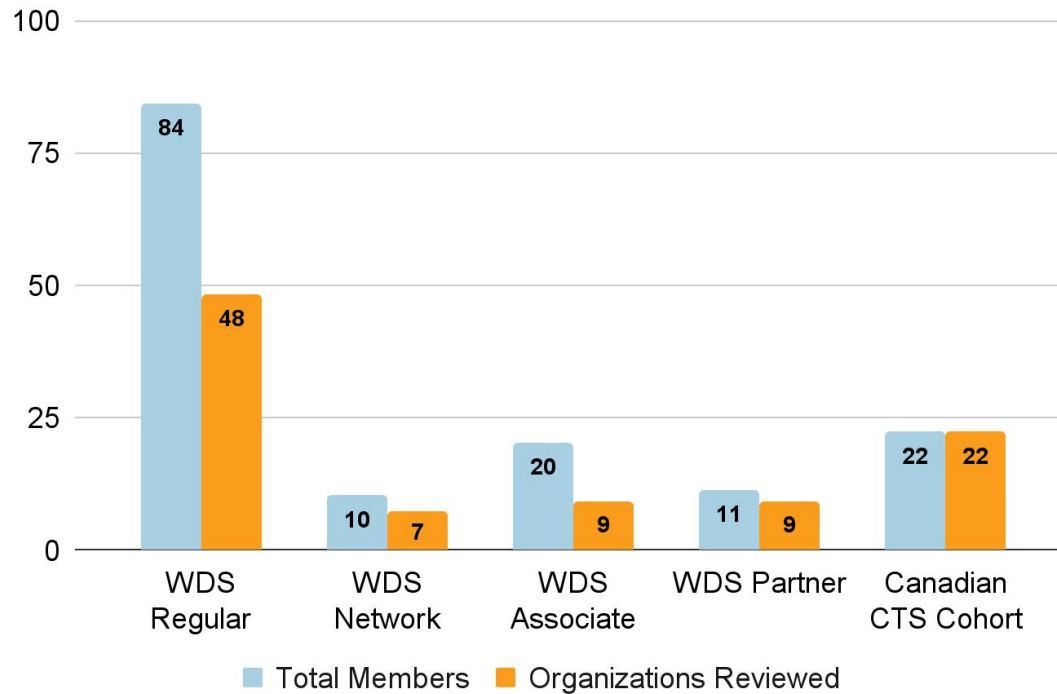
14. Mission: What is the Mission statement?

Qualitative

Your answer

Example of Reviewer Instrument

Overview



Organizations reviewed compared to total number of WDS and Canadian CTS Cohort members

Documents included within this assessment were all pulled from public-facing websites for WDS and Canadian CTS Cohort members. One document or source of information was reviewed for each organization for a total of:

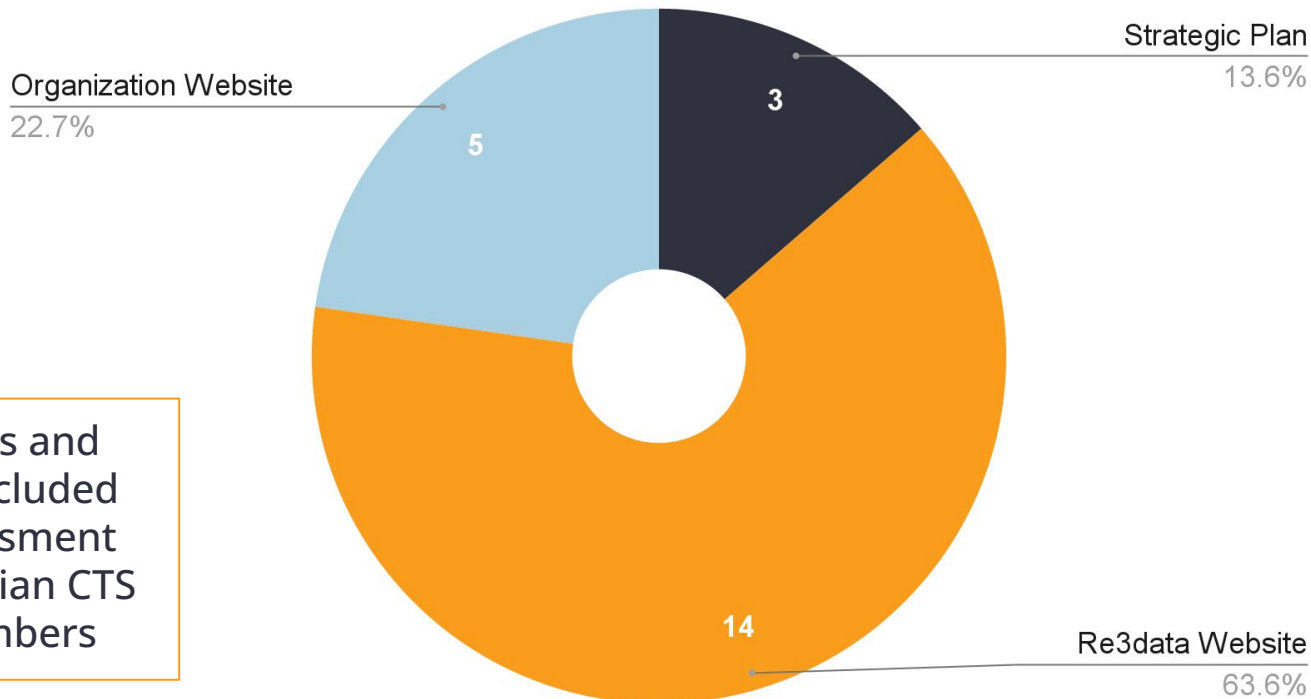
- 73/125 WDS Members
- 22/22 Canadian CTS Cohort Members

Overview

In total, **only 36 out of the 95 reviewed documents were a strategic plan or technical roadmap**. Therefore, documents such as CoreTrustSeal applications, annual reports, data policies, and journal articles were included.

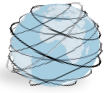
Additionally, websites were included as sources of information for the Canadian CTS Cohort due to difficulty locating additional documents.

Canadian CTS Cohort



Source types and numbers included in the assessment from Canadian CTS Cohort members

Overview



Source types and numbers included in the assessment from WDS members

WDS Members

Statutes

1.4%

Journal Article

5.5%

Data Policy

2.7%

Annual Report

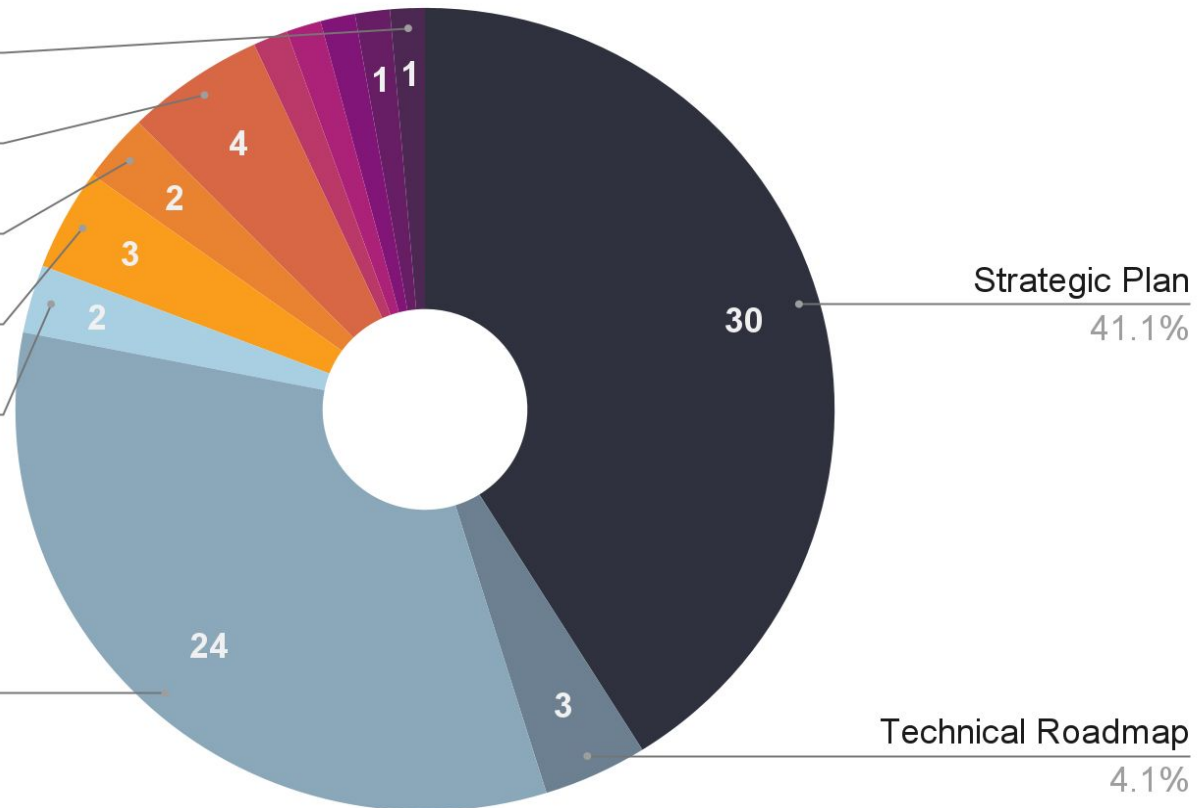
4.1%

Fact Sheet

2.7%

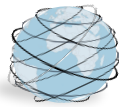
CoreTrustSeal

32.9%

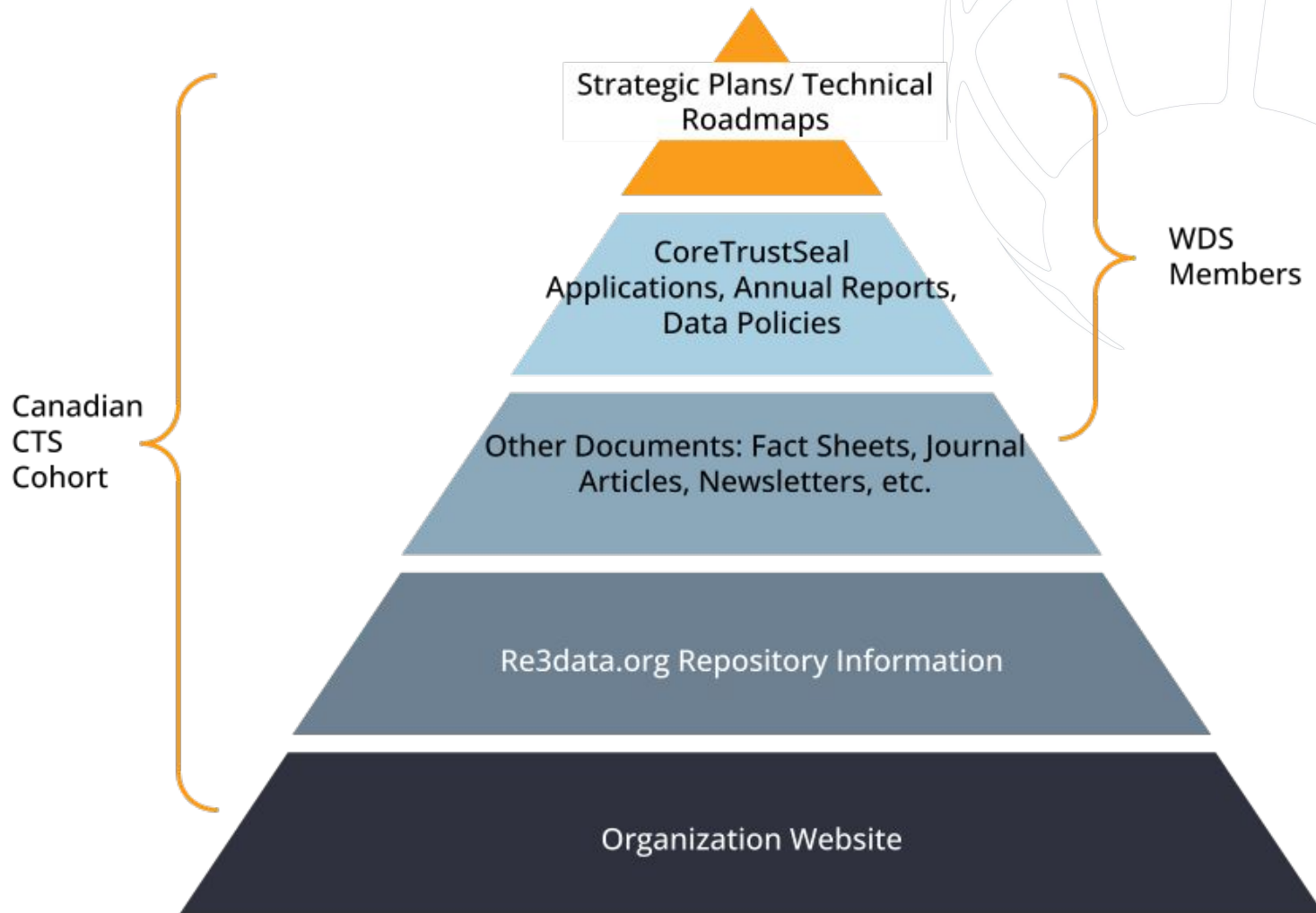


Values not shown: Other Report: 1 (1.4%), Technical Reference Manual: 1 (1.4%), Newsletter: 1 (1.4%), Magazine: 1 (1.4%)

Overview



Hierarchy of Document Priority



Results



Summary



All Organizations Reviewed

Language

German

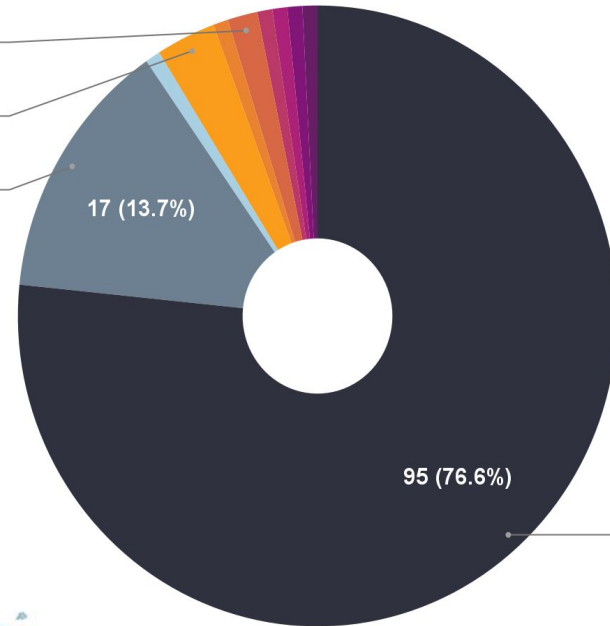
1.6%

Chinese

3.2%

French

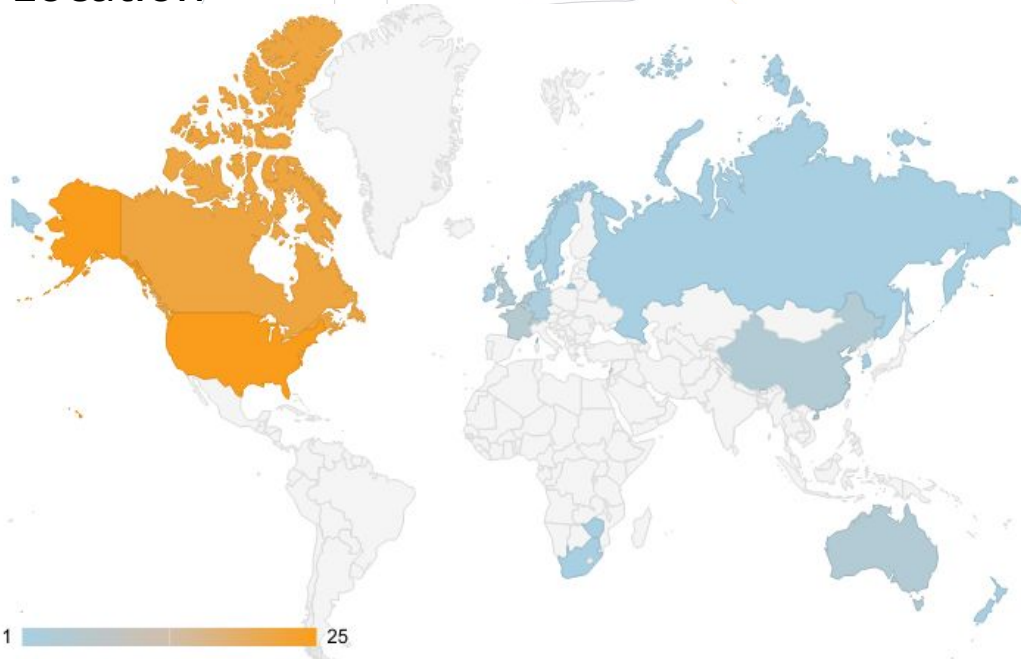
13.7%



English
76.6%

Values not shown: Swedish: 1 (0.8%), Chinese: 4 (3.2%), Norwegian: 1 (0.8%), German: 2 (1.6%), Russian: 1 (0.8%), Dutch: 1 (0.8%), Arabic: 1 (0.8%), Korean: 1 (0.8%)

Location

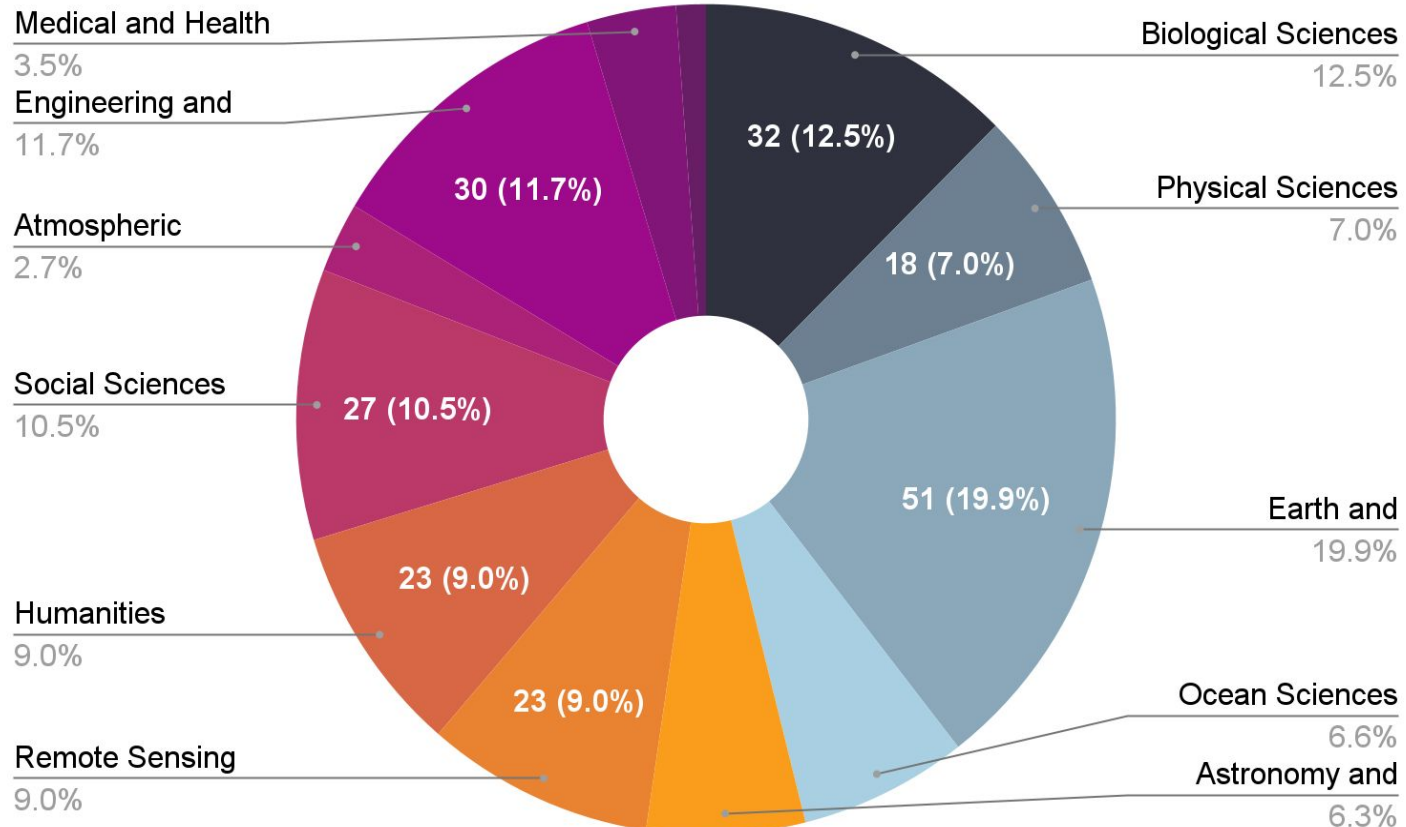


Summary



All Organizations Reviewed

Domains

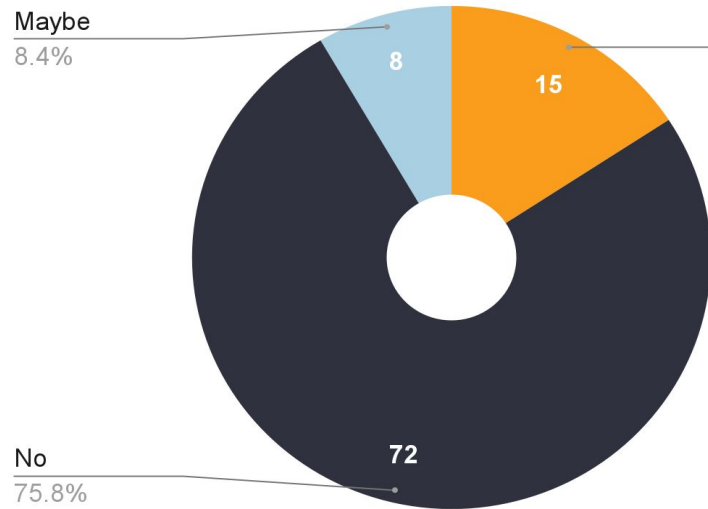


Values not shown: Ocean Sciences: 17 (6.6%), Astronomy and Space Sciences: 16 (6.3%), Atmospheric Sciences: 7 (2.7%), Medical and Health Sciences: 9 (3.5%), Mathematics: 3 (1.2%)

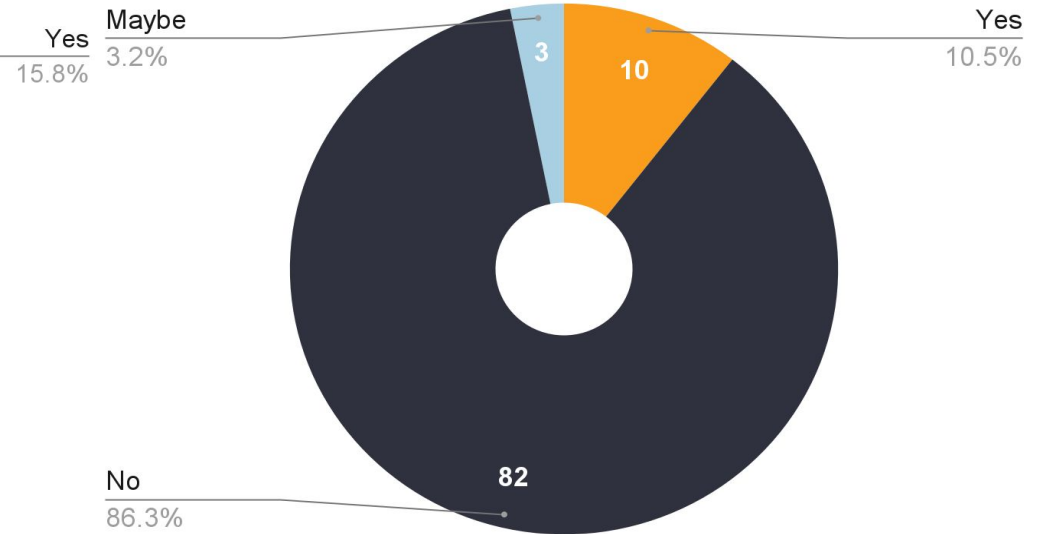
Summary

All Organizations Reviewed

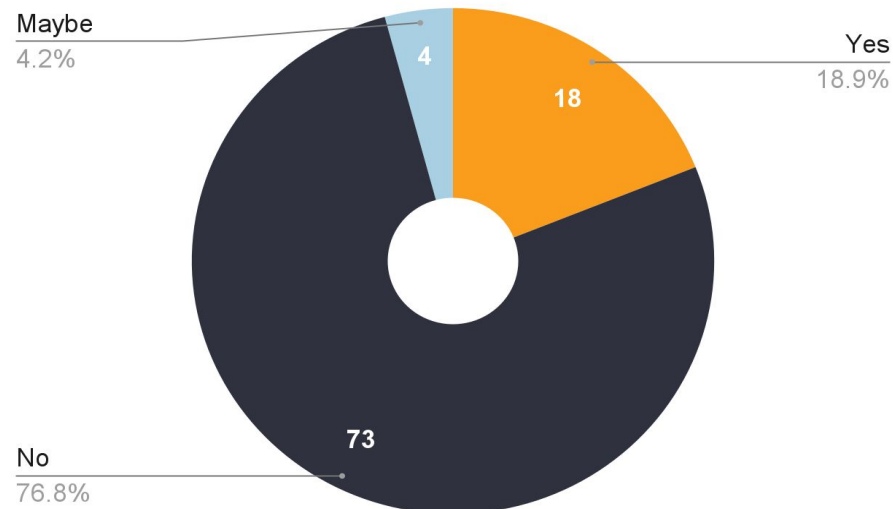
References to Polar Activities



References to Indigenous Groups



References to Advancing Diversity



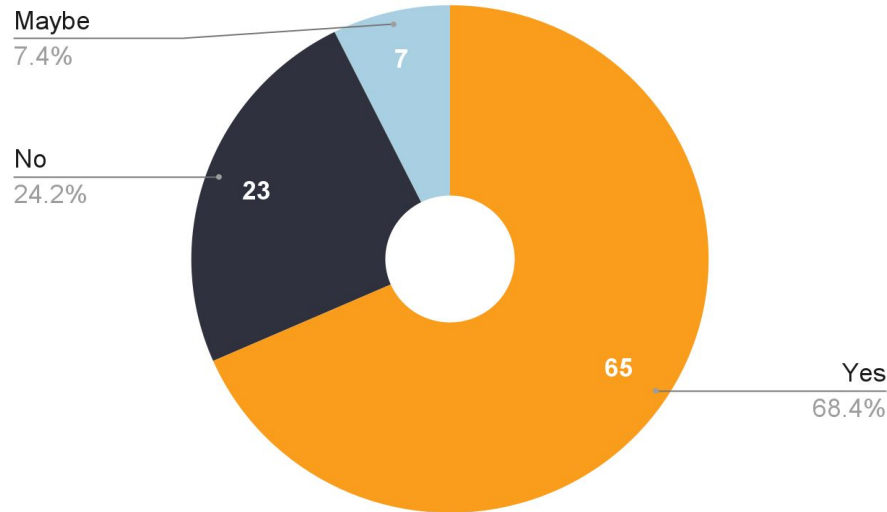
Note: Maybe indicates references that may indicate mention or involvement in these areas but could not be definitively confirmed in the document reviewed.

References to Mission, Vision and Guiding Principles

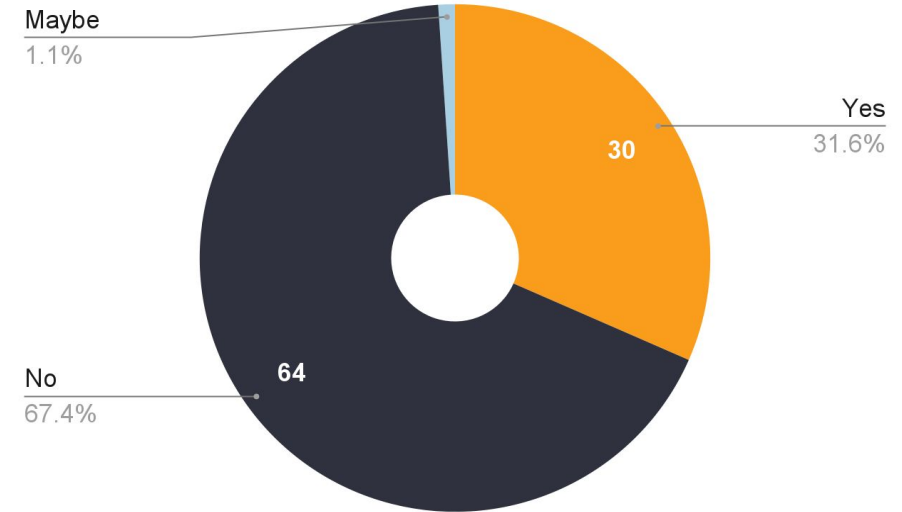


All Organizations
Reviewed

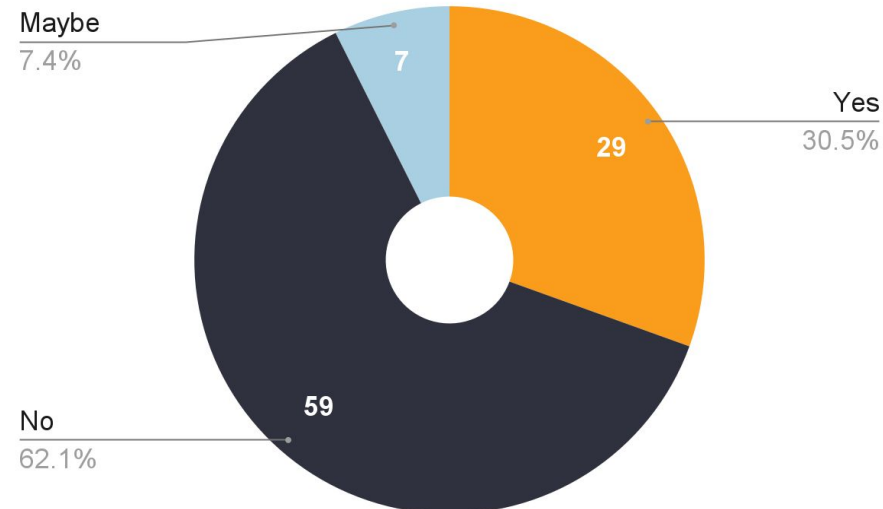
Mission Statement



Vision Statement

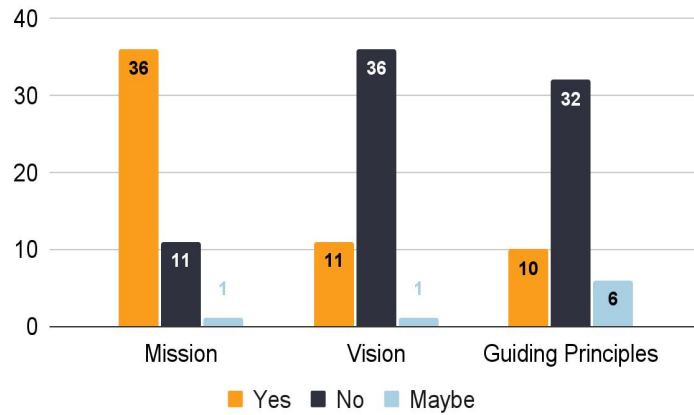


Guiding Principles

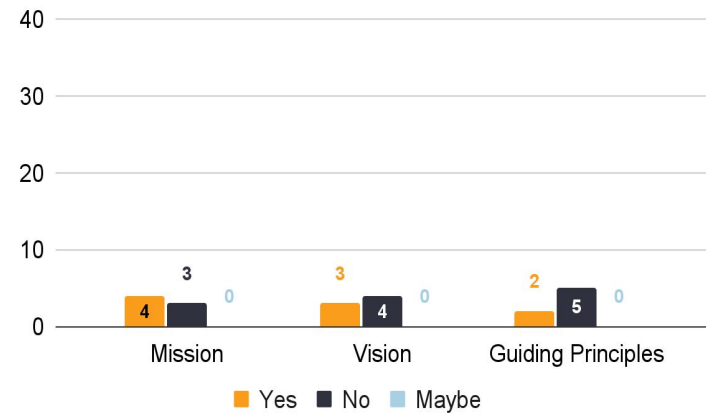


Note: Maybe indicates organizations that may have reference to a mission or vision statement or guiding principles but could not be definitively confirmed in the document reviewed.

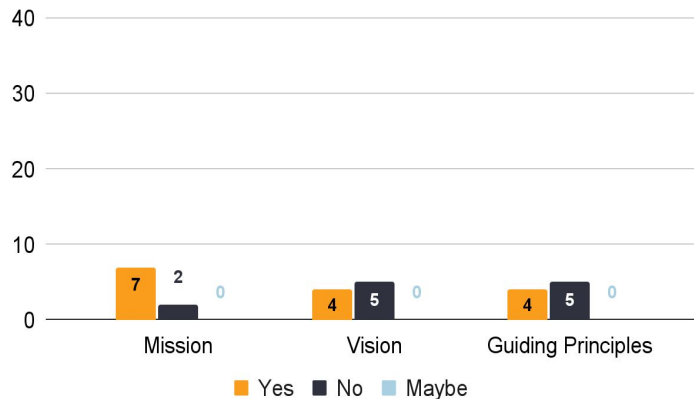
WDS Regular



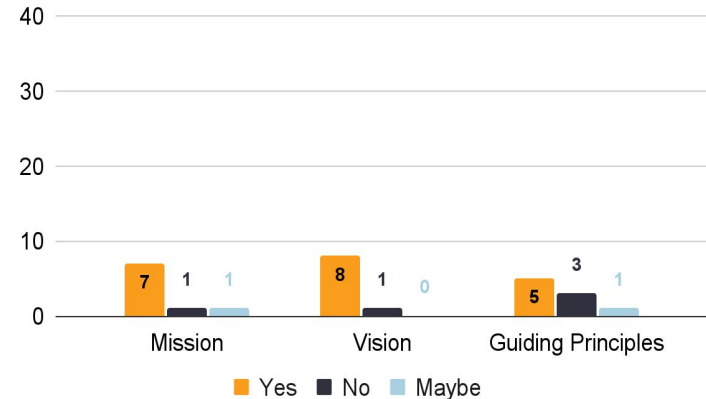
WDS Network



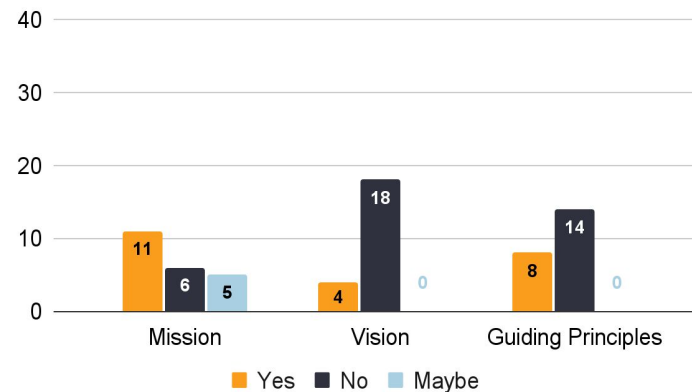
WDS Associate



WDS Partner



Canadian CTS Cohort



**By Membership
Type**

WDS Regular

- Advance research
- Provide high-quality data products
- Increase collaboration with partners and users
- Ensure preservation of data
- Protect the environment and address global change

WDS Associate

- Promote open data
- Form international communities
- Advance diversity and inclusivity

WDS Network

- Create infrastructure that provides high-quality data products
- Advocate for open data

WDS Partner

- Advance user engagement and build stakeholder relationships
- Support international cooperation
- Support the needs of data users

Canadian CTS Cohort

- Promote knowledge and data sharing with regards to the FAIR principles
- Build and maintain strong partnerships and ensure community engagement
- Advance diversity and inclusivity

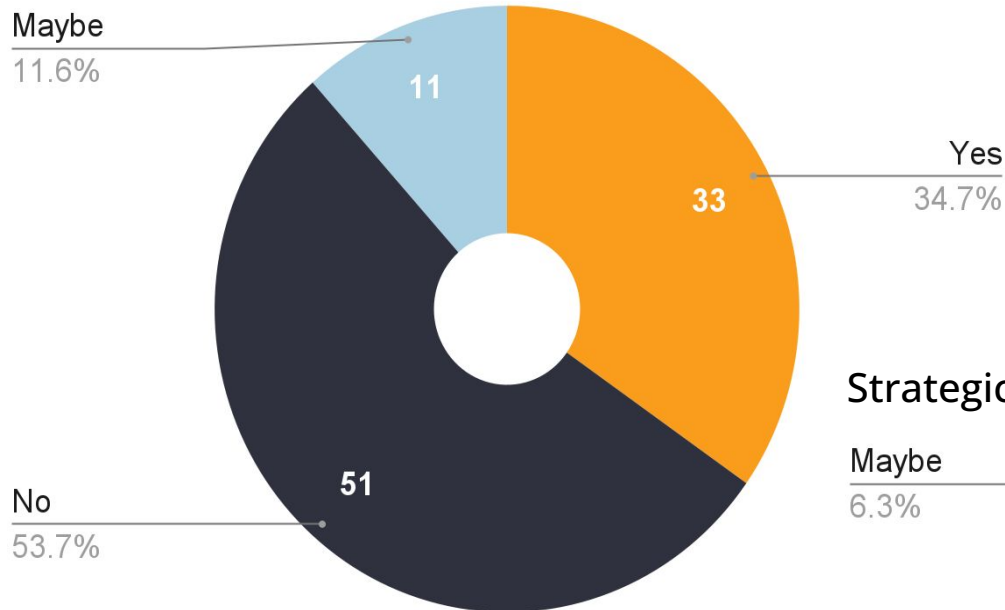
**By Membership
Type**

References to Long-Term Goals and Strategic Priorities

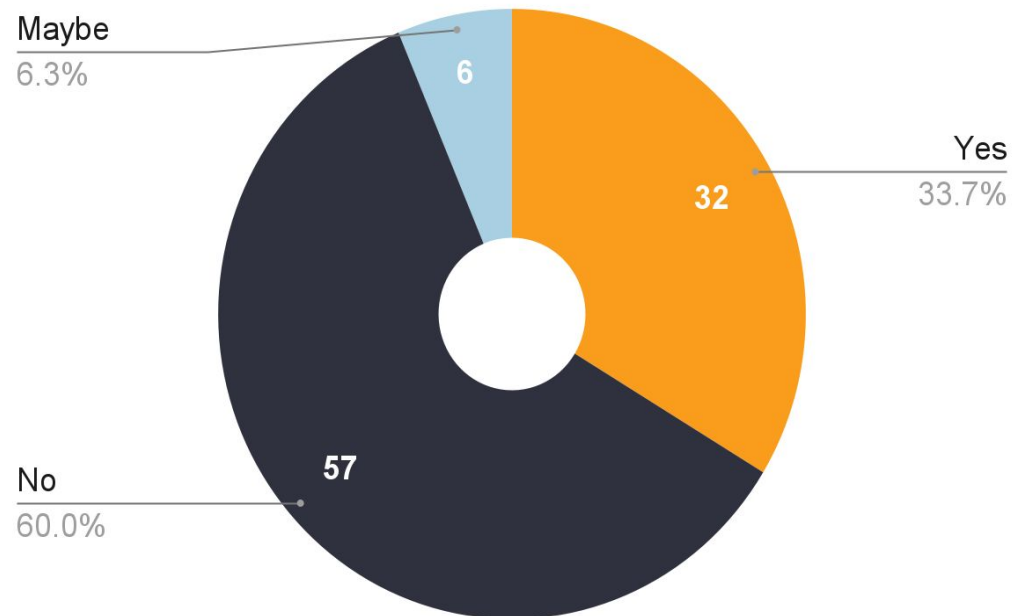


All Organizations
Reviewed

Long-Term Goals

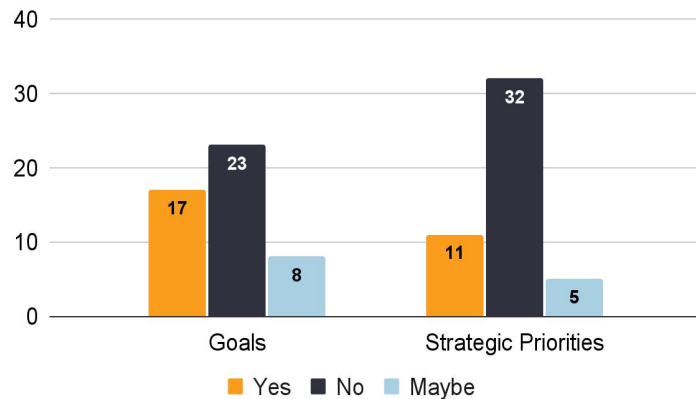


Strategic Priorities

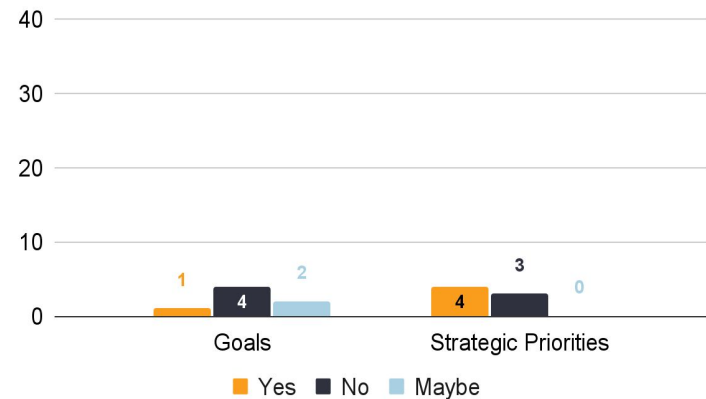


Note: Maybe indicates organizations that may have reference to goals or strategic priorities but could not be definitively confirmed in the document reviewed.

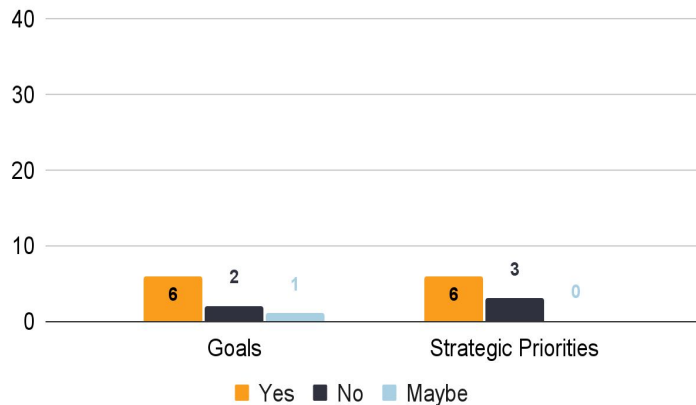
WDS Regular



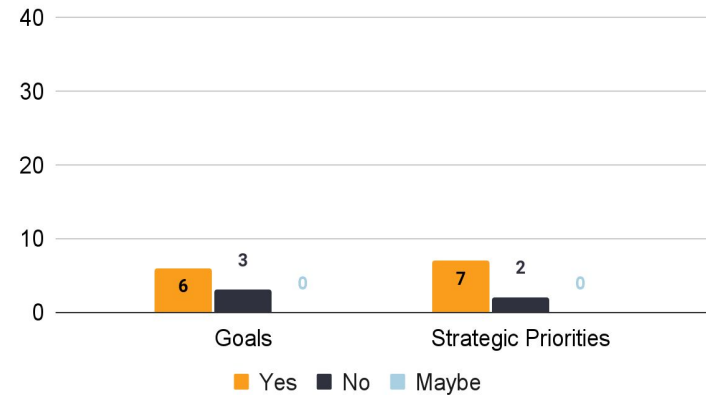
WDS Network



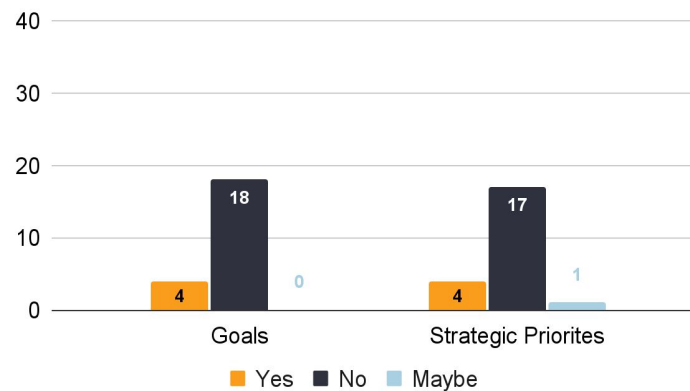
WDS Associate



WDS Partner



Canadian CTS Cohort



**By Membership
Type**

WDS Regular

- Provide high-quality data products and services
- Ensure long-term data preservation
- Conduct impactful research
- Develop and maintain partnerships locally and internationally

WDS Associate

- Ensure financial sustainability
- Create products and services that impact the user
- Attract funders and partnerships, improve outreach and facilitate engagement

WDS Network

- Improve data collection facilities
- Prioritize user experience
- Integrate with other scientific communities

WDS Partner

- Strengthen and promote the role and impact of the organization
- Provide research and services that aim to solve global problems
- Advance infrastructure, address data gaps and improve data curation

Canadian CTS Cohort

- Increase engagement with stakeholders and partners
- Expand data holdings and improve infrastructure
- Ensure data collections are high-quality, reliable, and easy to access

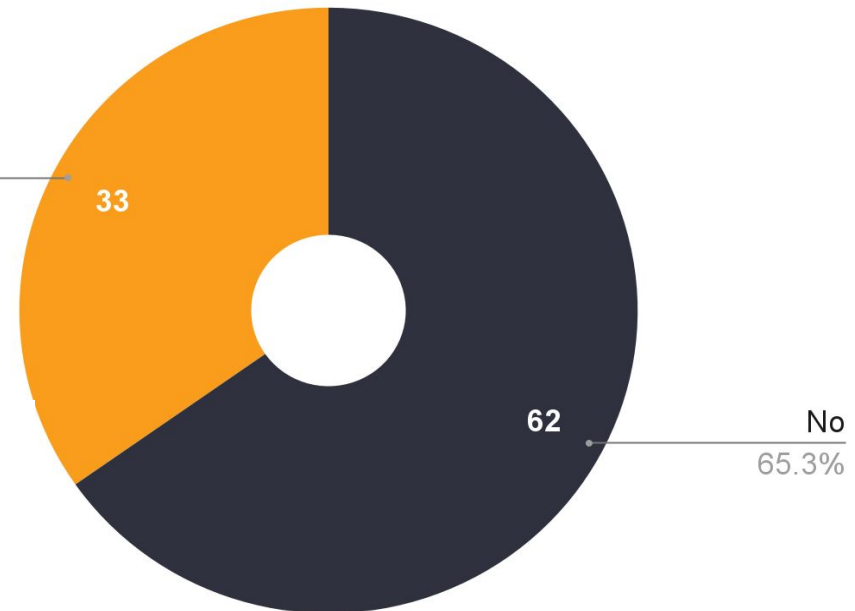
By Membership Type

References to Challenges and Technical Requirements

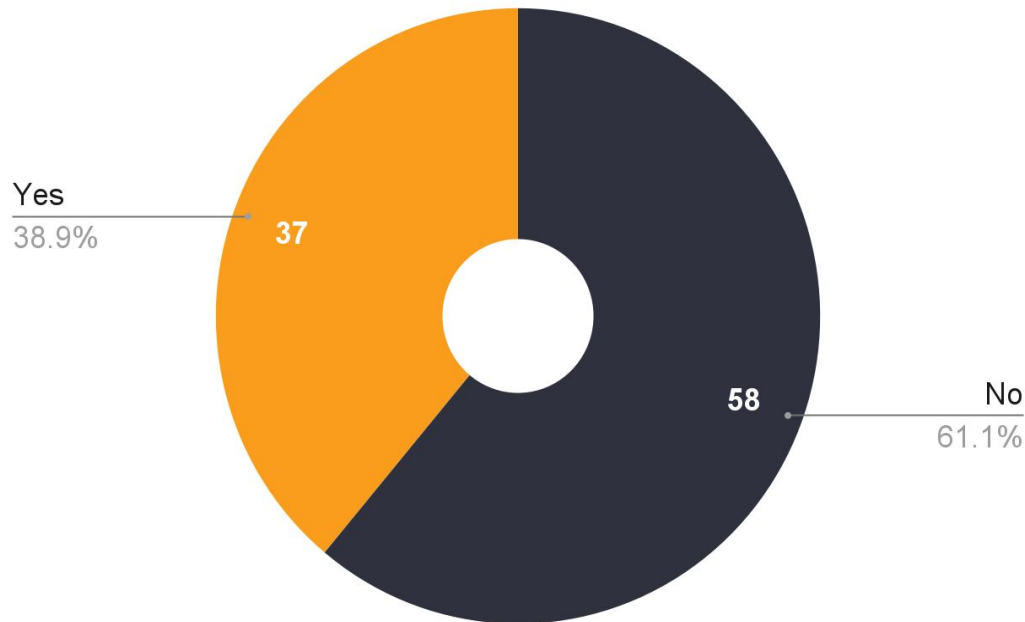


All Organizations
Reviewed

Technical Requirements

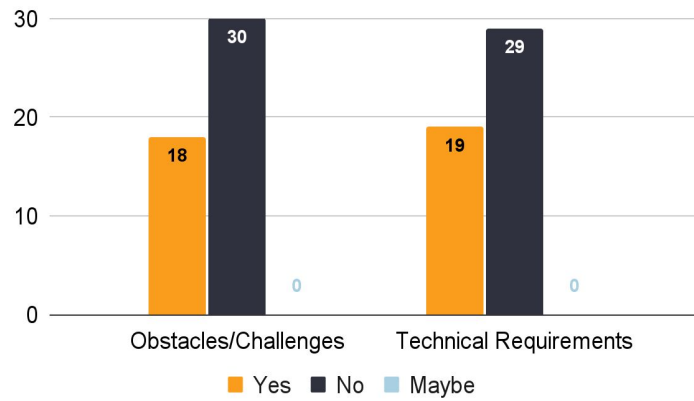


Challenges/ Obstacles

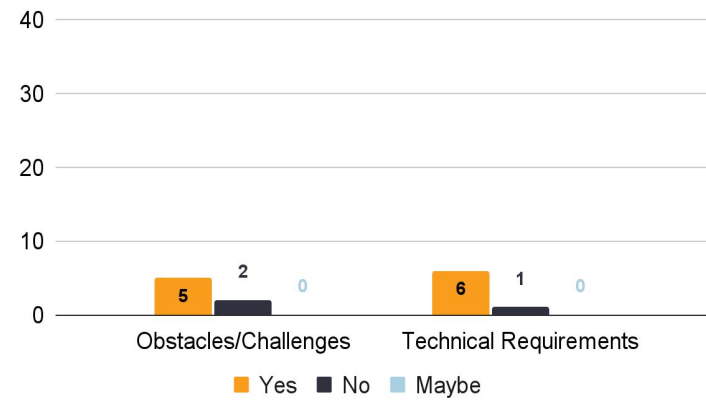


Note: Maybe indicates organizations that may have reference to obstacles/ challenges or technical requirements but could not be definitively confirmed in the document reviewed.

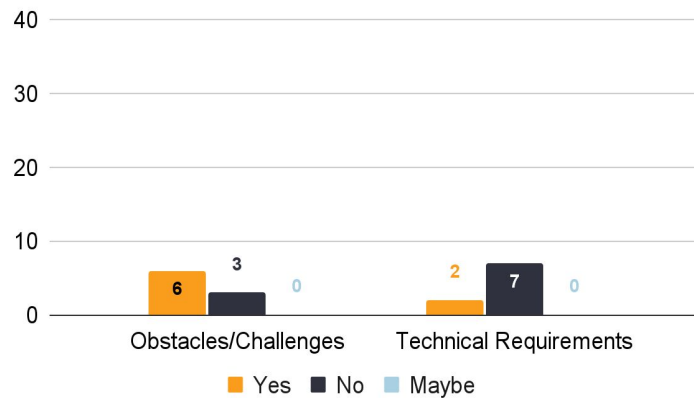
WDS Regular



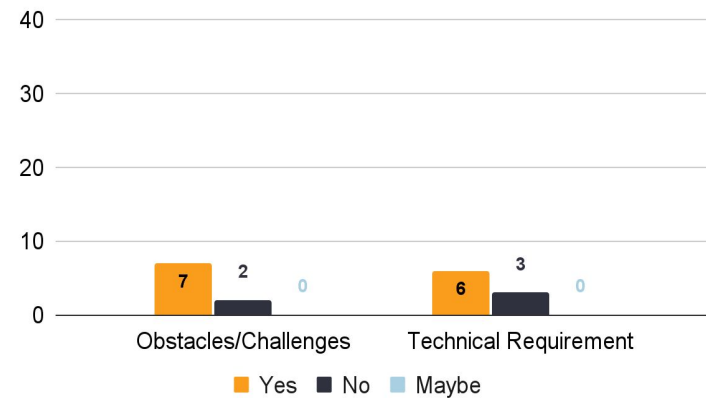
WDS Network



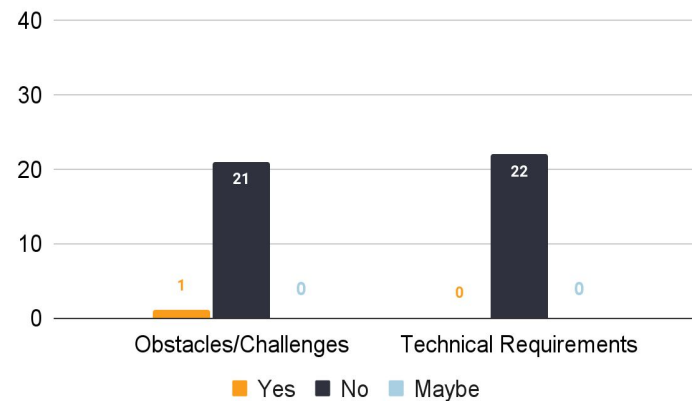
WDS Associate



WDS Partner



Canadian CTS Cohort



**By Membership
Type**

WDS Regular

- Lack of funding for continuous upward growth, funding for staff, and funding for improving infrastructure
- Aging infrastructure
- Increases in data volume beyond data storage capacity

WDS Network

- Meeting data specifications and quality with what is feasible economically
- Limitations of data transfer capacity
- Lack of data coverage in certain geographic areas

WDS Associate

- Need for consistent funding
- Challenges with only receiving project based funding

WDS Partner

- Need for long-term funding
- Challenges relating to data access, data sharing and reuse
- Need for infrastructure modernization

Canadian CTS Cohort

- Challenges with cybersecurity risks relating to technological change

**By Membership
Type**

Global Biodata Coalition (GBC) Criteria



Global Biodata Coalition (GBC) Criteria

The **GBC** is an initiative designed to support funders by identifying and prioritizing fundamental data resources that should be maintained as part of the worldwide life science infrastructure.

To properly explain the criteria the WDS and Canadian CTS Cohort the criteria have been split in three categories:

1. *Organization overview*
2. *Data storage, curation, protection and use*
3. *Partnerships, user support and outreach*

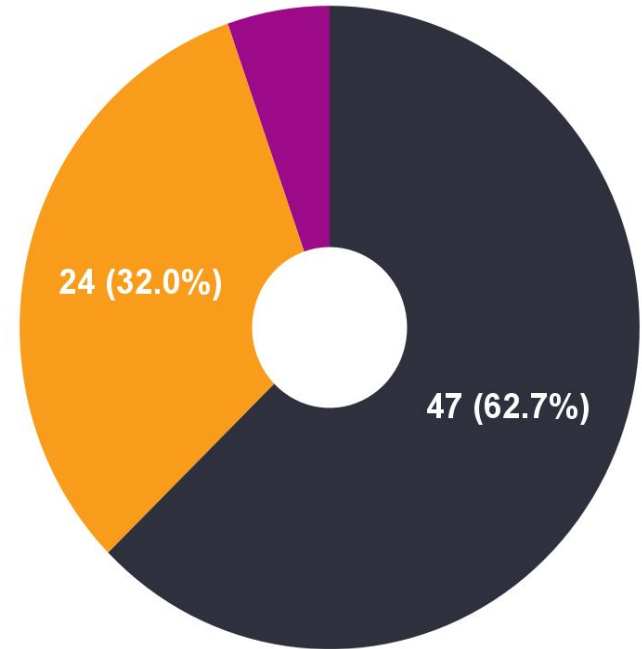
GBC - Organization Overview



All Organizations Reviewed

Host Organization Type

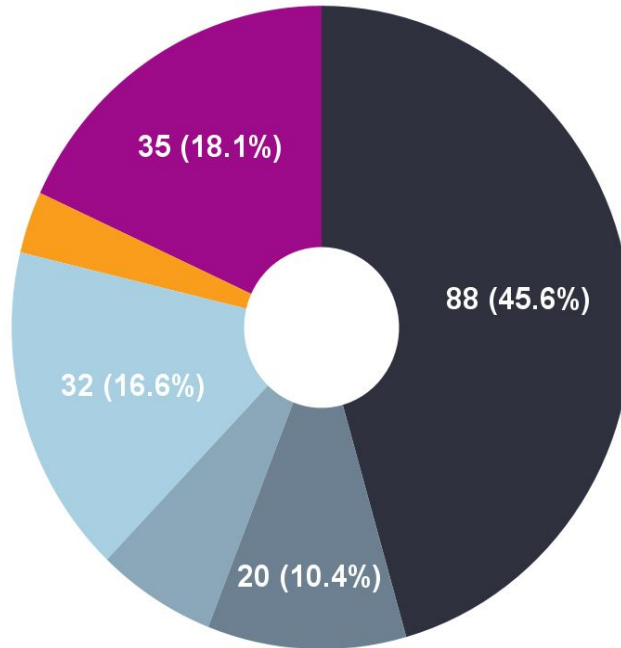
- Academic Institution
- Government Entity
- Non-Governmental Organization



Not shown: General Public/ Community: 12 (6.2%), Industry: 6 (3.1%)

Designated User Community

- Academic/ Research Users
- Employees
- General Public/ Community
- Government/ Policy Makers
- Industry
- Partners



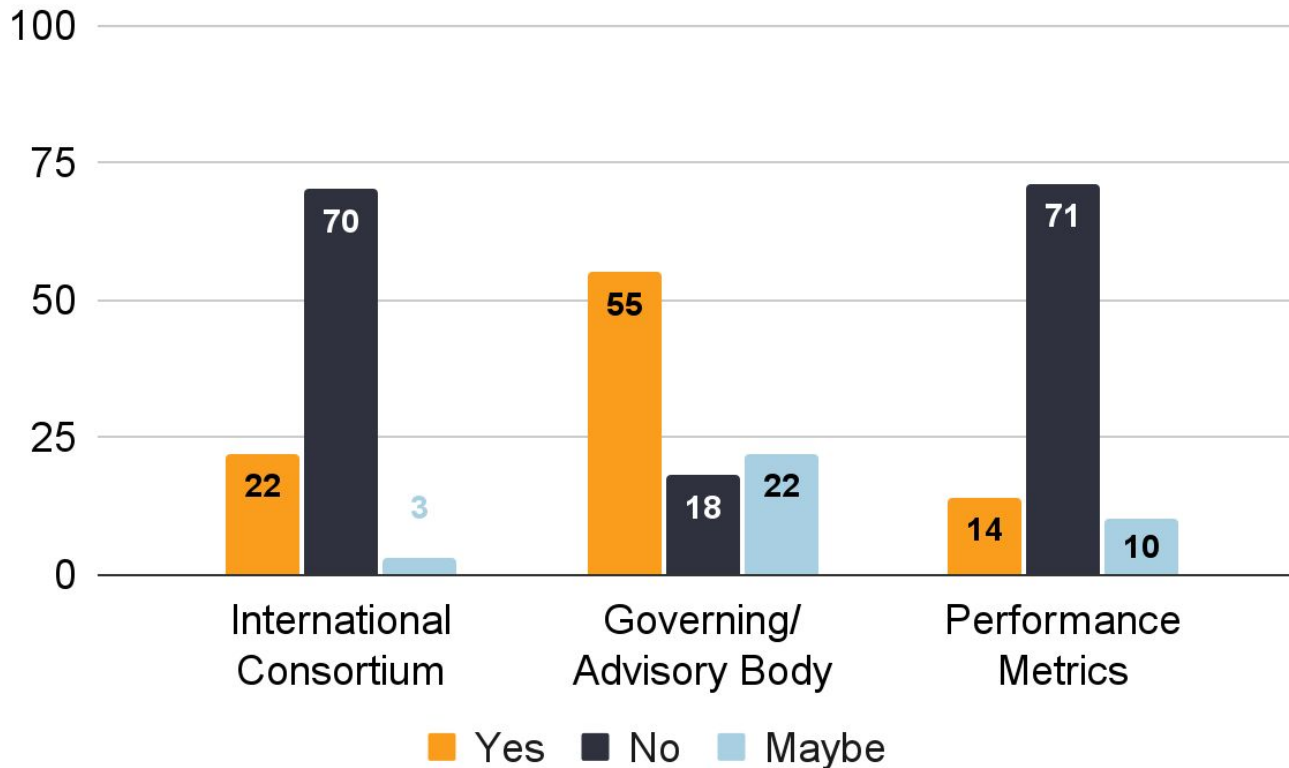
Not shown: Non-Governmental Organization: 4 (5.3%)

GBC - Organization Overview



All Organizations Reviewed

References to Being Part of an International Consortium, Having a Governing/ Advisory Body or Performance Metrics



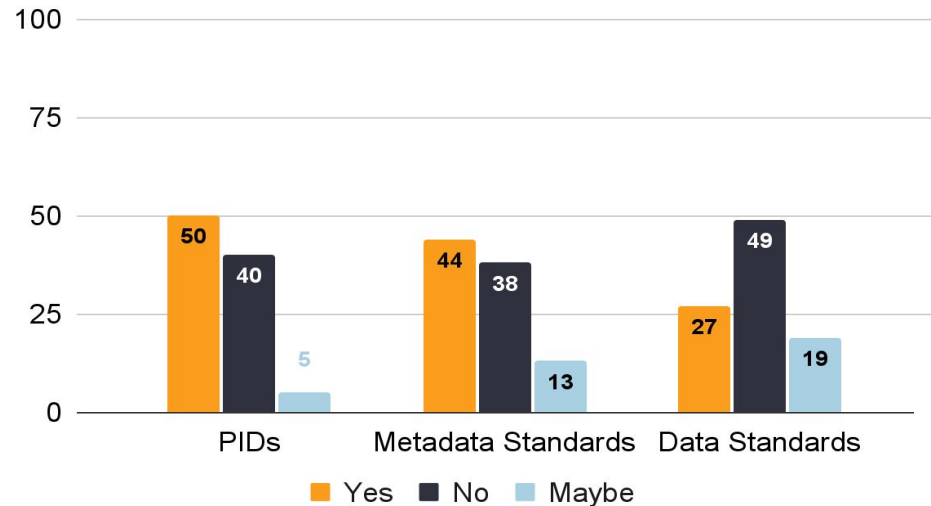
Note: Maybe indicates organizations that may have reference to being part of an international consortium, having a governing/advisory body or referencing performance metrics but could not be definitively confirmed in the document reviewed.

GBC - Data Storage, Curation and Protection



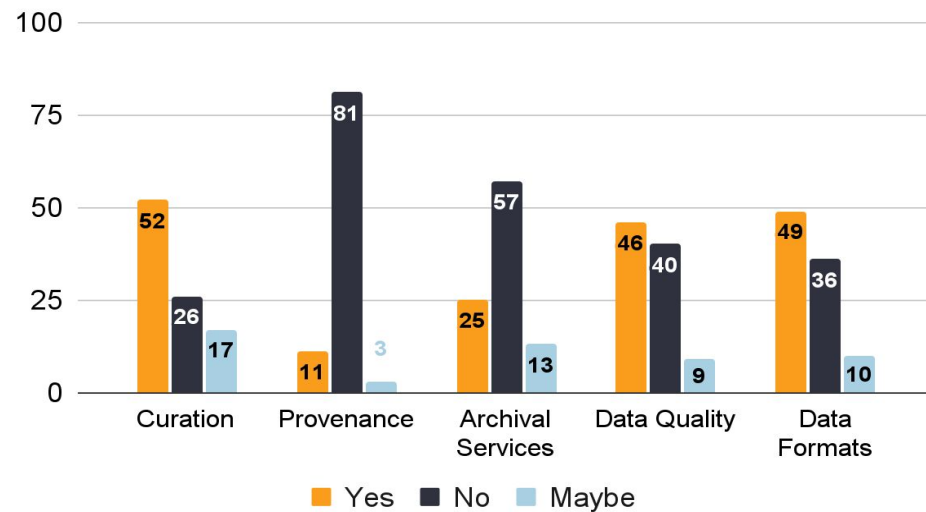
All Organizations
Reviewed

References to the Use of PIDs, Metadata Standards or Data Standards



Note: Maybe indicates organizations that may have reference to the use of PIDs, metadata and/or data standards but could not be definitively confirmed in the document reviewed.

References to Data Curation, the Ability to Provide Provenance to Data Holdings, Archival Services, Data Quality or Data Formats



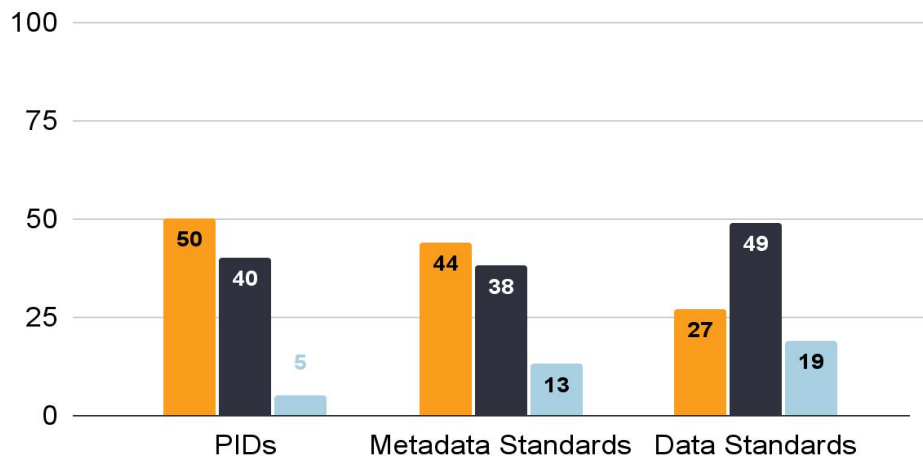
Note: Maybe indicates organizations that may have reference to data curation, ability to provide provenance to data holdings, archival services, data quality and/or data formats but could not be definitively confirmed in the document reviewed.

GBC - Data Storage, Curation and Protection

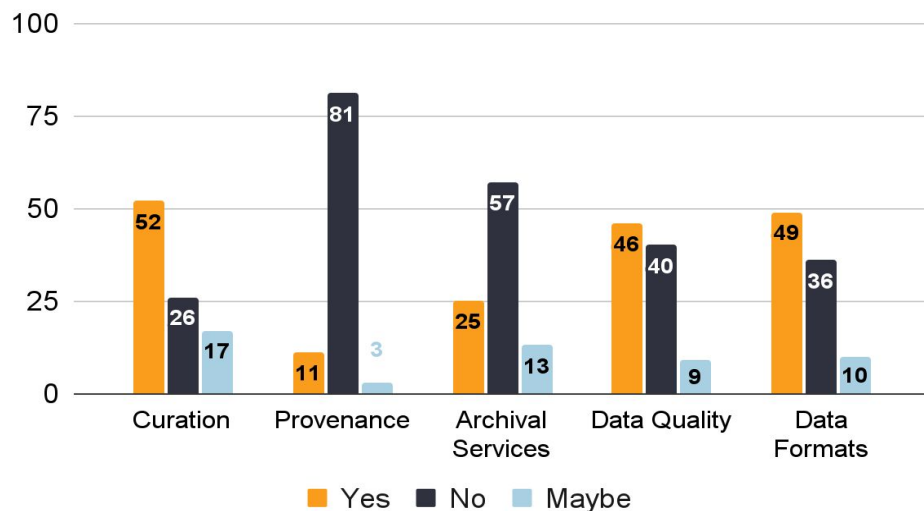


All Organizations Reviewed

References to the Use of PIDs, Metadata Standards or Data Standards



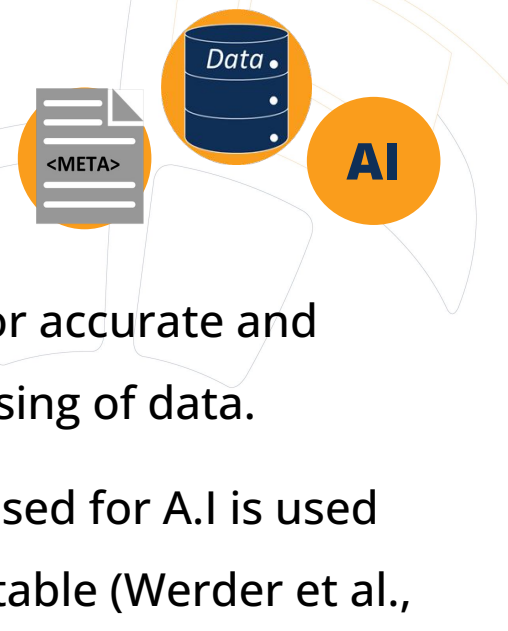
References to Data Curation, the Ability to Provide Provenance to Data Holdings, Archival Services, Data Quality or Data Formats



A list of specific standards, policies and practices that have been mentioned by WDS and Canadian CTS Cohort members has been created

Note: Maybe indicates organizations that may have reference to data curation, ability to provide provenance to data holdings, archival services, data quality and/or data formats but could not be definitively confirmed in the document reviewed.

Data Provenance



The rise of Artificial Intelligence (A.I) demonstrates the need for accurate and transparent information surrounding the creation and processing of data.

Comprehensive provenance information ensures that data used for A.I is used responsibly, and A.I use remains fair, transparent and accountable (Werder et al., 2022)

“The repository should provide evidence to show that it operates a data and metadata management system that maintains provenance information to ensure authenticity from deposit, and through curation and preservation to the point of access.”

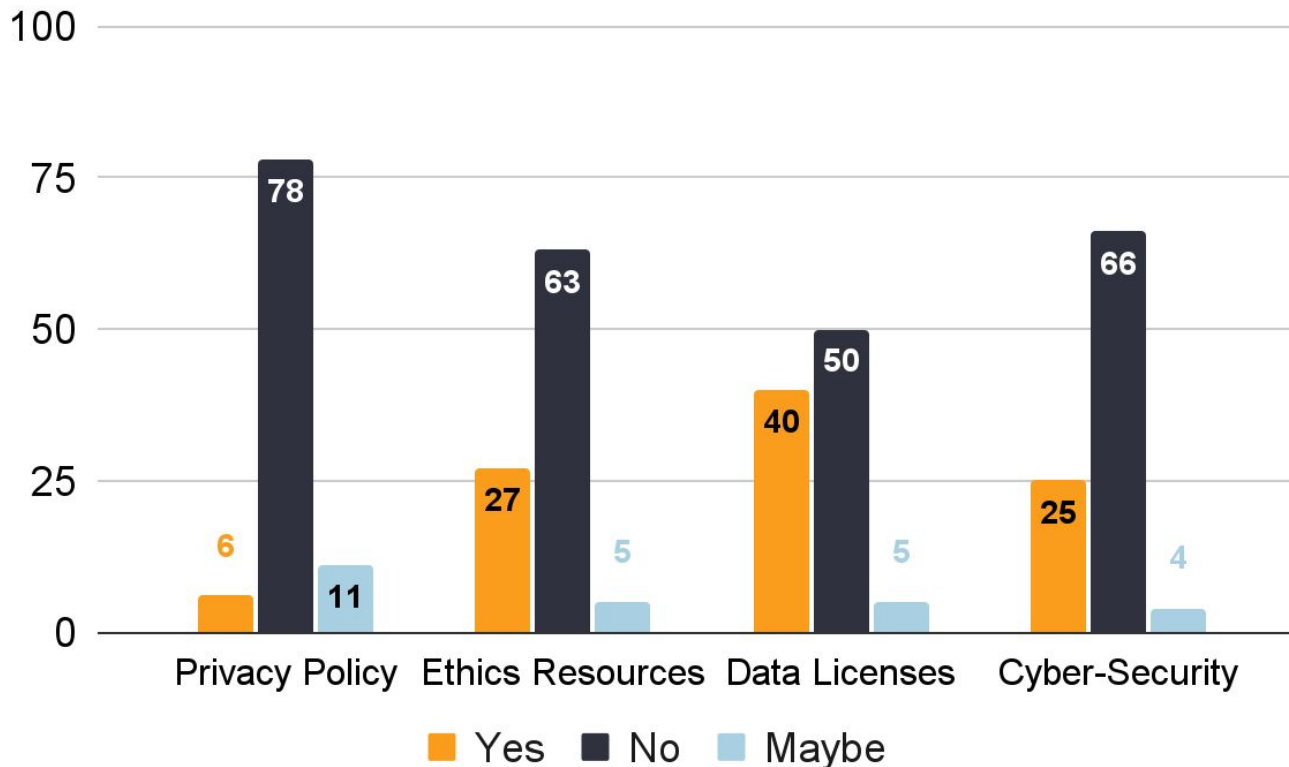
– CoreTrustSeal Requirements 2023 - 2025

GBC - Data Storage, Curation and Protection



All Organizations Reviewed

References to the Use of a Privacy Policy, Ethics Resources, Data Licenses or Cyber-Security Measures



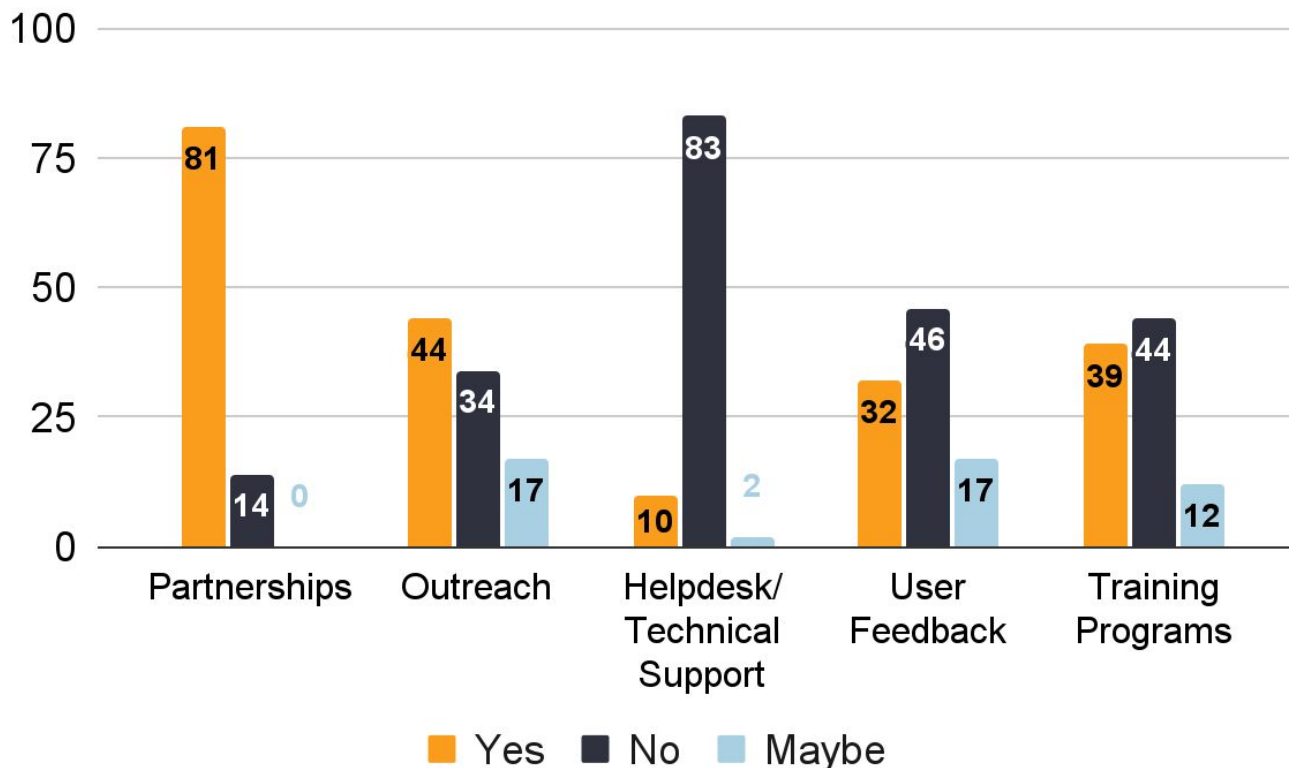
Note: Maybe indicates organizations that may have referenced the use of a privacy policy, ethics resources, data licenses or cyber-security measures but could not be definitively confirmed in the document reviewed.

GBC - Partnerships, User Support and Outreach



All Organizations
Reviewed

References to Current Partnerships, Community Outreach and User Engagement, Helpdesk or Technical Support, User Feedback, or Training Programs



Note: Maybe indicates organizations that may have reference to current partnerships, mechanisms for community outreach and user engagement, the ability to provide helpdesk or technical support, mechanisms for user feedback, or training programs but could not be definitively confirmed in the document reviewed.

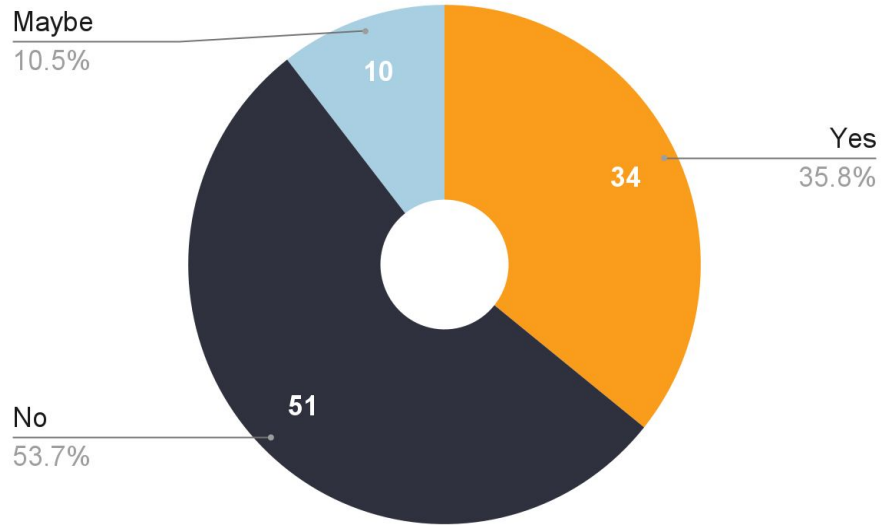
Key Points



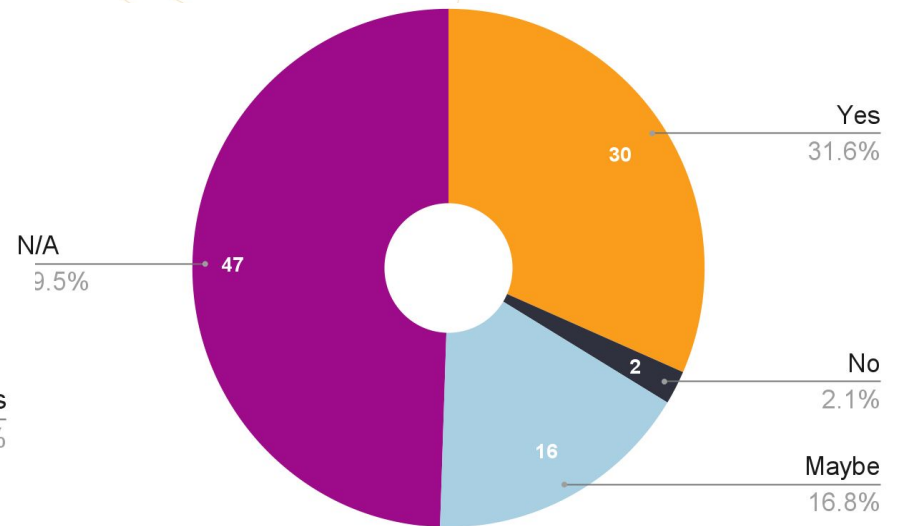
All Organizations Reviewed

The two largest concerns arising from our review across all member types were sustainability and creating effective partnerships with others entities in the research data space.

References to Planning/Outreach for New Partnerships



References to Sustainable and Long-Term Funding



Note: N/A indicates organizations that did not reference funding or sustainability in the document reviewed.

Moving Forward



Based on the analysis from this assessment, the **WDS recommends** assisting Canadian repositories and WDS members in several ways. These recommended actions are based on the [World Data System Action Plan 2022 through 2024](#).

- **Help repositories create a strategic plan or technical roadmap to outline how they achieving their goals**
- **Launch a webinar series on overcoming common challenges**
- **Create resource areas on the WDS website for ways repositories may communicate how they are implementing TRUST, FAIR, and other data standards**
- **Communicate funding opportunities to members**

It is the recommendation of the research team that WDS commit to an ongoing needs assessment for data repositories to be compiled on an annual basis.

Thank You

This work is supported by the Digital Research Alliance of Canada and Ocean Networks Canada in partnership with the World Data System International Program Office hosted by the University of Tennessee Oak Ridge Innovation Institute, supported by a cooperative agreement (DE-SC0021915) with the U.S. Department of Energy Office of Science (PI Suzie Allard).



CC BY license

Research Analysis: WDS Members and Canadian CTS Cohort Needs Assessment
ITO - March 14th/ 2023

ito-director@oceannetworks.ca
wds-ito.org



References

- CoreTrustSeal Standards and Certification Board. (2022). CoreTrustSeal Requirements 2023-2025 (V01.00). Zenodo. <https://doi.org/10.5281/zenodo.7051012>
- Digital Research Alliance of Canada. (2023). CoreTrustSeal Certification Support Cohort & Funding. <https://alliancecan.ca/en/coretrustseal-certification-support-cohort-funding>
- Global Biodata Coalition. (2022). Global Core Biodata Resources: Concept and Selection Process. <https://doi.org/10.5281/zenodo.5845116>
- WDS. (2022). Action Plan 2022 - 2024. World Data System. <https://worlddatasystem.org/about/action-plan/>
- Werder, K., Ramesh, B., Zhang, R. (2022). Establishing data provenance for Responsible Artificial Intelligence Systems. ACM Transactions on Management Information Systems, 13(2), 1-23. <https://doi.org/10.1145/3503488>