

REACHING FOR INTERSTELLAR

Data and information
sharing at CSA

CONDENSED REPORT

**POWERED
BY
DATA**

Leslie L. Cheung, PhD & Michael Lenczner

March 2018

About Powered by Data

Powered by Data's mission is to maximize the availability and impact of data for public good. Through an approach that blends data policy and data strategy development, Powered by Data helps establish infrastructure and governance frameworks that will enable the social sector to better share, use, and learn from data. Powered by Data works with nonprofits and civil society groups, government, funders, and global data initiatives.

Powered by Data operates on Tides Canada's shared platform, which supports on-the-ground efforts to create uncommon solutions for the common good.

To learn more, visit our website at: www.poweredbydata.org

POWERED
BY
DATA

Authors

Leslie L. Cheung, PhD
Partnership Development Director & Open Data Lead, Powered by Data

Michael Lenczner
Director, Powered by Data

Contributors

Catherine Doldirina, PhD
Independant Consultant, and member of the International Institute of Space Law

Layout

Lorraine Chuen
Communications Lead, Powered by Data

TABLE OF CONTENTS

Overview	4
Part 1:	
Comparing data and information sharing practices of space agencies	8
Part 2:	
Feedback from consultation with potential users of CSA data and information. .	14
Part 3:	
Conclusions and recommendations	27



OVERVIEW

The Canadian Space Agency (CSA) works to advance the knowledge of space through science and to ensure that space, science, and technology provide social and economic benefits for Canadians. It has a mandate to comply with and fulfil the Treasury Board Secretariat's (TBS) Directive on Open Government. Under this mandate, CSA must organise and release its data and information of business value to support transparency, accountability, citizen engagement, and benefits from re-use of these data and information. The CSA Open Government team would like to develop an open government plan that both fulfils its open government obligations and furthers the delivery of CSA's broader mandate.

The Open Government team at CSA hired Powered by Data to provide strategic advice and guidance on CSA's open data and information publishing practices, based on our experience and expertise in open data publication by government institutions and organisations. The recommendations and guidance are based on best practices of similar government institutions, potential user/stakeholder needs, and open government trends.

This report is divided into three sections described below:

1. Comparing data and information sharing practices of space agencies
2. Feedback from consultation with potential users of CSA data and information
3. Conclusions and recommendations

1. Comparing data and information sharing practices of space agencies

The first section of this report reviews and compares the **current data and information sharing activities of six other space agencies against those practices of CSA**. We reviewed publicly available documents and websites of six space agencies; NASA (USA), ESA (Intergovernmental), CNES (France), DLR (Germany), ASI (Italy), and JAXA (Japan). We present a summary and comparison table of some of the best data and information policies and practices of these six space agencies compared to CSA. The section provides guidance on how to improve CSA's data and information sharing practices based on common practices from other space agencies and relevant open data trends.

The largest barrier to opening more scientific data and information at CSA is that CSA does not currently own or hold much of the scientific data from projects to which it contributes due to the standard policy and practice of CSA funding through grants and contributions, and contracts. Despite this, there is a history and an appetite at CSA and among CSA's third-party researchers and partners to open CSA-funded science data and information. While CSA has a government-wide open-data policy, they are lacking a data management framework that would provide CSA employees across the organisation clarity on their priorities and vision of the organisation's goals in relation to what data can achieve for CSA. This would provide an operating framework across the organisation to establish priorities, create processes, encourage collaboration, allocate resources, and make decisions about how to collect, manage, and archive data and information.

If this framework supports open by default policies throughout the organisation, this helps space agencies achieve the goals that are inherent in open government initiatives: Not only transparency, but also increased innovation, collaboration, economic development, engagement, and evidence-based decision-making in the space and science sector. CSA is unable to maximise the potential value of data and information that have access and use restrictions. Furthermore, a centralised agency site to access data and information, providing access to third-party data and information, and active user engagement help increase the discoverability and use of data and information of CSA-funded data and information.

2. Feedback from consultations with potential users of CSA data and information

This section presents the **findings of a two-stage consultation with potential space data and information users**. The consultations provided concrete examples of potential use and requirements of CSA data and information.

The first stage of the consultation was an online questionnaire distributed to potential data user groups or individuals via social media outreach or email. Follow-up interviews were conducted with a subset of respondents to gather more details and document the specific needs of potential users from a variety of sectors and skills levels.

Consultation findings highlighted a range of applications for space data and information across sectors, such as in policy, education, and technology development. Consultation feedback also sheds light on the data and information requirements requested by potential users to facilitate their access and use of space data and information, such as the use of specific licensing, data formats, and the discoverability and searchability of CSA data assets.

Providing additional services to meet the needs of a potential users with different levels of capacity, such as providing multiple products for the same asset, building tools to view and analyse data, or clearing the rights over data not owned or held by CSA, will be dependant on the resources available within CSA. Establishing the priorities of CSA to provide these additional services would be facilitated by an articulated data management policy.

3. Conclusions and recommendations

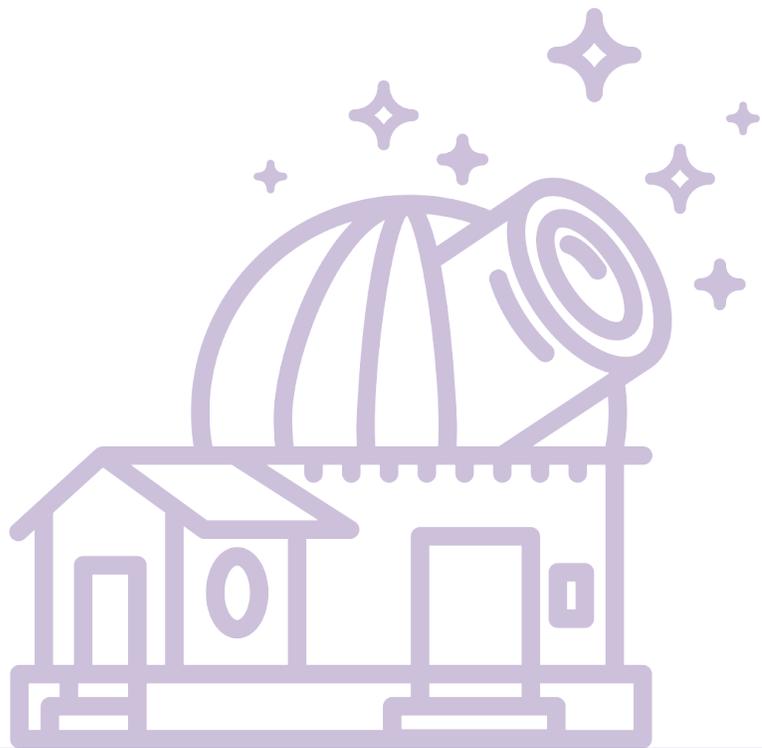
Both the review across space agencies and the user consultations indicate **a policy gap between what CSA is mandated to do via its Open Government obligations and how it would like to better accomplish the agency's broader mission by opening more CSA-funded science data**. Lacking an agency-wide data management policy, the CSA Open Government team is unable to reconcile these competing priorities.

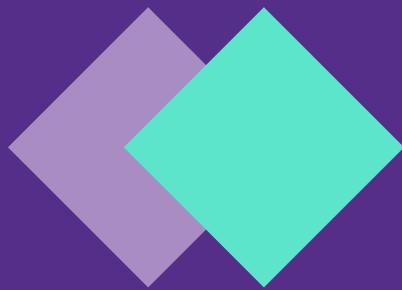
We propose three scenarios that require an increasing amount of collaboration with internal and external partners, and an increasing amount of resources. However, the more collaboration and resources required, the higher the level of potential impact of open data and information to enable innovation, economic development, collaboration, and evidence-based decision-making in the sector.

The first option, "focus on Open Government obligations" has the Open Government team using open government to accomplish its mission by concentrating its efforts on data and information CSA owns. The second option, "build on the status quo", involves the CSA Open Government team continuing and improving the work it is currently doing: knowledge dissemination for both CSA owned assets, but also those scientific data and information to which it contributes, but does not own. The third option, "Agency-wide data and information management framework" involves CSA providing a framework for data and information sharing that would provide the

mandate for any change to policies and practice throughout the organisation.

These three scenarios are not distinct options but exist on a scale of possibilities that could be adopted in a non consecutive order. Furthermore, the recommendations are not mutually exclusive to one scenario, but could be implemented cumulatively. Each scenario is presented with advantages and disadvantages to help CSA consider the consequences of each. Recommendations for each scenario are intended to help maximise the impact of each option. The report will conclude with some considerations about the level of organisational support and resources that are available at CSA.





PART 1

COMPARING DATA AND INFORMATION SHARING PRACTICES OF SPACE AGENCIES



METHODOLOGY

To compare the data and information publishing practices at CSA to other space agencies, we reviewed the practices of six space agencies and those of CSA. The other space agencies reviewed are: NASA (USA), ESA (Intergovernmental organisation), CNES (France), DLR (Germany), ASI (Italy), and JAXA (Japan). These six space agencies were selected because of their relative advanced maturity for publishing data and information. In addition, for most of the agencies selected, a legal obligation to publish data and other content is imposed by legislative or regulatory acts, as is in the case of CSA. Finally, these six agencies have close ties to CSA due to a high number of joint projects and activities.

Information was collected from a variety of sources available through each agency's website to gain a best-effort portrait of the data and information publishing practices at each space agency. We did not restrict ourselves to only open data sharing practices, but rather kept a broad scope to better understand what data and information is publicly available and to what extent each agency publishes open data and information.

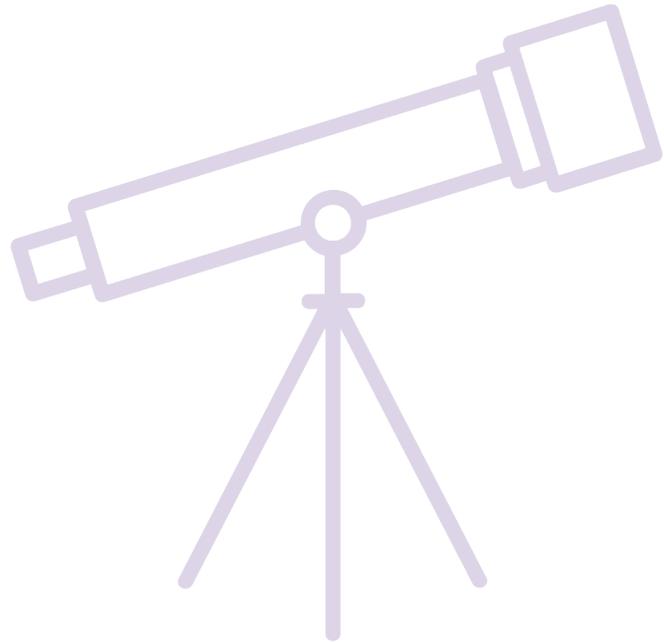
With the exception of CSA, the limited scope of the desktop reviews meant that the data sources were restricted to publicly available information only. Therefore, the level of detail available for each agency will differ. This includes, for example, policies and practices around publishing data, and knowledge dissemination activities. We were also limited by our language abilities. All of the agencies had English websites, but some information may only be available in the predominant language of the agency's home country. This is particularly true for JAXA where much of its information is in Japanese.

Information about CSA was supplemented with a review of a number of internal documents, including, but not limited to reports written by staff or external contractors, data management planning documents, and statistics from TBS on open asset downloads. In addition, we interviewed CSA employees from the three program sectors, as well as those with knowledge of contracts, and grants and contribution agreements with third-parties. These interviews helped us understand some of the barriers that CSA faces and its success stories of opening science-based data and information.

In total, we interviewed 20 CSA employees including:

- The Access to Information and Open Data Coordinator and their two staff members;
- Two employees involved in contracting;
- One employee who works on the management grants and contribution agreements across the organisation;
- 14 employees from the three program sectors; Space Utilisation, Space Exploration, and Space Science & Technology.¹

Below is a summary and comparison table of the sharing practices of CSA compared to six other space agencies.



¹ The Open Government Coordinator at CSA provided a list of 15 staff members from the program sectors. Thirteen responded and agreed to be interviewed, plus an additional name was added upon suggestion from another interviewee.



FINDINGS

Open data is growing in popularity amidst a quickly changing digital environment. Germany and France have adopted open data policies over the last two years, following in the footsteps of governments like Canada and the US. As more and more governmental organisations adopt these types of policies, the opportunities to achieve the intention and vision of these policies increase, as better ways to regulate and implement effective practices emerge.

NASA has been relatively successful in opening large amounts of data and related products. NASA has an agency-wide data management plan that is modelled on an open science framework. This allows NASA to place conditions on its funding that require its funded science activities to operate openly by default. NASA archives resulting data and makes it available under an open license. There are acceptable exemptions to this policy, such as security concerns or temporary embargos on data to allow time for researchers to publish. The reasons for these exemptions need to be stipulated in a data management plan before receiving funding.

An agency-wide data management policy is not the norm among the space agencies reviewed, but there are other successful open by default initiatives. Specifically Sentinel products from Copernicus, the European Union's Earth observation programme which operates open by default. Policies and initiatives like these, have helped facilitate the sharing of open data that are created within participating space agencies and by third-parties. They provide an operating framework across the organisation to establish priorities, create processes, encourage collaboration, allocate resources, and make decisions about how to collect, manage, and archive data and information. This helps these space agencies achieve the goals that are inherent in open government initiatives: Not only transparency, but also innovation, collaboration, economic development, engagement, and evidence-based decision-making in the space and science sector. Space agencies are unable to maximise the potential value of data and information that are under access and use restrictions, as is the case with large quantities of data and information at CSA and all of the agencies reviewed, with the exception of NASA.

A centralised agency website can increase discoverability of data and information assets, yet only two of the space agencies (NASA and ESA) reviewed had centralised sites.

France and Canada's national governments operate open government portals, but CNES and CSA each provide additional data and information on their agency websites that exceed the assets permissible on the portal, as some of them may not be openly licensed or held by the agency. But neither of these space agencies have a centralised portal on their websites to access these assets, like ESA or NASA do. However, all of the space agencies reviewed provide access to data on third-party websites for projects to which the agency has contributed funds. By providing information about these sometimes hard-to-find third-party websites, this further increases the discoverability of space-agency funded data and information by potential users.

Finally, each of the space agencies reviewed engage with potential users in a variety of ways. Social media, press releases, and blogs are some of the ways these space agencies inform potential users about their data and information assets. Hackathons and events like conferences or workshops are deeper engagement activities that expose potential users to agency data and information. Finally, agency websites are used as engagement tools for specific audiences. Age appropriate resources for children and youth are common.

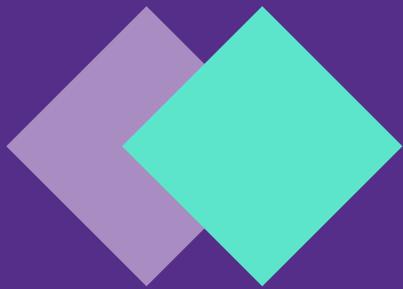
Below is a summary table of the common characteristics of the reviewed space agencies that facilitate publication of, access to, and use of space data and information.

This section looked to six other space agencies to better understand the common and successful data and information sharing practices to compare to the practices of CSA. The next section will summarise the findings of a consultation with potential space and data information users to better understand their needs. The final part of this report will provide three distinct scenarios for CSA to consider as it moves forward in its data and information sharing work.

COMPARING DATA AND INFORMATION SHARING PRACTICES OF SEVEN SPACE AGENCIES

	<u>CSA</u>	<u>NASA</u>	<u>ESA</u>	<u>CNES</u>	<u>DLR</u>	<u>ASI</u>	<u>JAXA</u>
Country or Region	Canada	USA	Inter-governmental	France	Germany	Italy	Japan
Government-wide open data policy or regulation	✓	✓	✗	✓	✓	✓	✗
Agency-wide data management framework	✗	✓	✗	✗	✗	✗	✗
Agency-funded third-party data and information open by default (with valid exemptions)	○	✓	○	○	○	○	✗
Data and information released free of restrictions on access and use	○	✓	○	○	○	○	○
Centralised agency site to access data and information assets	✗	✓	✓	✗	✗	✗	✗
Provides access to data on external third-party websites	✓	✓	✓	✓	✓	✓	✓
Active user engagement	✓	✓	✓	✓	○	✓	✓

LEGEND ✓ Yes ✗ No ○ Sometimes



PART 2

FEEDBACK FROM CONSULTATION WITH POTENTIAL USERS OF CSA DATA AND INFORMATION



METHODOLOGY

The Open Government team at CSA wanted to learn more about the types of potential users of space data and information and their needs, beyond that of the academic research community. To help construct a list of potential users, Powered by Data created outreach target groups by focusing on the audiences that CSA defines on its website: educators, industry, media, museums and science centres, and the scientific community. CSA program sector staff were asked to contribute individuals and organisations to the list based on their knowledge of the sector focusing on those with which they do not usually consult. Powered by Data added to this list based on our knowledge of the open data community and our past experiences with open data consultations. The final list included 116 individuals or organisations, with almost all of them located in Canada.

Outreach methods began with either a targeted social media message via Twitter or by email, with an explicit request for those contacts to share the questionnaire widely with interested parties, including employees or members of their organisations and networks. Email reminders or telephone calls to specific groups were employed to further promote the questionnaire.

The online questionnaire consisted of eight questions and was available in both French and English. The questionnaire was available from January 31st, 2018 to February 9th, 2018 (10 days) and in that time, we received 235 responses².

These responses were analysed to provide a broad picture of the types of respondents and their responses. The content of the questionnaire was then analysed by purpose of use and the data requirements that would facilitate engagement with space data and information to draw out themes for each sector.

The second stage of this consultation were semi-structured interviews to detail use cases of space data and information. Interview subjects were selected from the 235 questionnaire respondents. To narrow the list of potential interviewees, all of those

² At close of the survey, there were 240 responses, however five responses were removed during analysis. One as a spoiled entry and four as duplicates of a previous response most likely due to submission errors.

respondents who did not provide contact information were eliminated, and then the participants were sorted by sector. From each category, respondents who provided more precise and specific use cases in their responses were selected for a longlist. This longlist was shortened by eliminating similar use cases and emphasising potential users located in Canada.

Requests for a 30-minute telephone interview were sent by email or made by telephone to 23 questionnaire respondents. In most cases interviews were one-on-one, except in two cases in which two respondents from the same organisation were interviewed together. In total, 12 individuals responded and were interviewed, representing 10 organisations.

During these interviews, participants were asked what data and information they would potentially use and for what purpose. Follow-up questions were asked regarding preferred format, periodicity, licensing, capacity to process data, and discoverability. Notes were recorded by the interviewer. The goal of these interviews was to better understand the needs of potential users who stated specific use cases. Content from these interviews provide concrete examples of use of space data and information.





FINDINGS

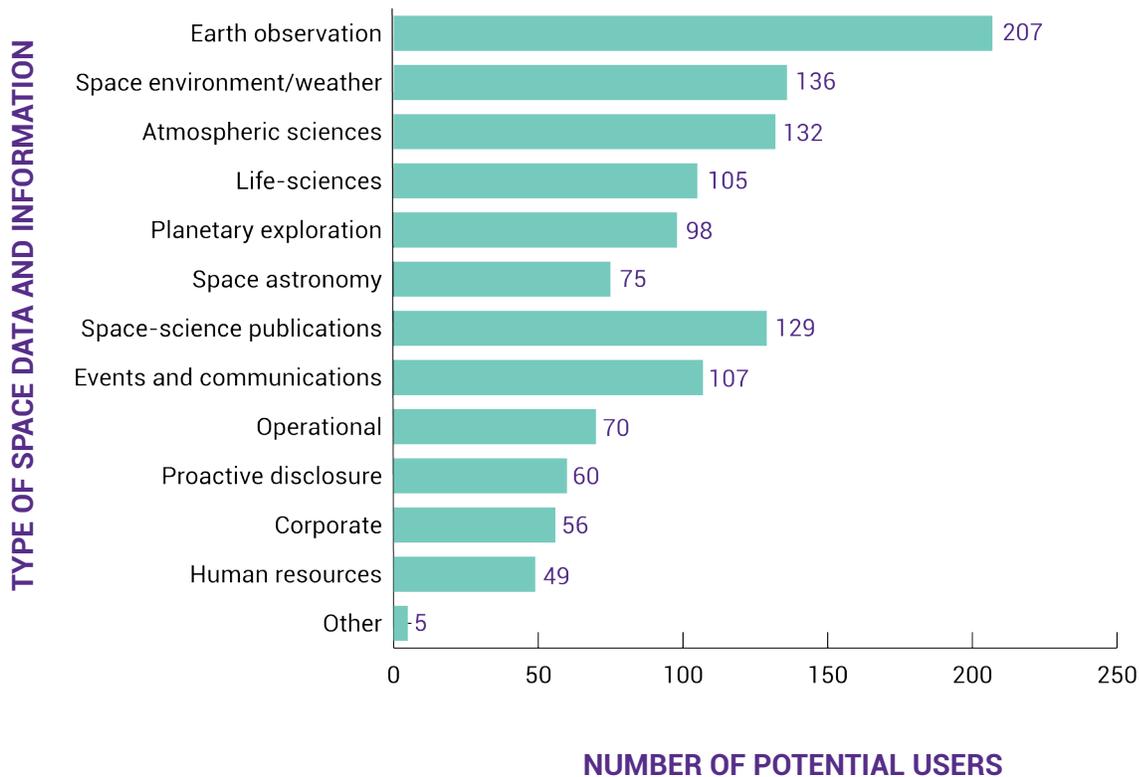
Out of 235 responses, over 88% (209) were submitted in English and over 11% (26) were submitted in French. While the outreach list only included groups and organisations in Canada, we used social media and requests for wide distribution to promote the questionnaire. As a result, at least half (117) of the respondents identified themselves as from abroad and at least 31% (73) of respondents were from Canada. These results should consider that the location of another 45 respondents could not be determined as they did not provide enough identifying information, such as a telephone number or a name of their affiliated organization.

Overall, from the sample of questionnaire respondents, there is more of an interest in scientific data and information rather than information related to the other functions of CSA. As for the research fields of scientific data available at CSA, Earth observation data and information were the top category of interest with 207 out of 235 respondents indicating an interest in accessing this type of scientific data (Figure 1). Space environment or space weather data, and atmospheric sciences data were of interest to 136 and 132 respondents, respectively. Life sciences was a scientific field that interested 105 respondents, followed by planetary exploration with 98 respondents indicating interest, and space astronomy data and information that was an interest of 75 respondents.

In terms of data and information related to different functions of the organisation, space science publications, such as peer-reviewed journals were desired by 129 respondents, followed by 107 respondents who were interested in data and information concerning events and communications at the CSA. The latter category includes items such as media releases, public opinion research, and educational material. Operational data, such as briefing materials and presentations were of interest to 70 respondents while 60 respondents were interested in data and information regarding proactive disclosure on such things as grants and contributions, contracts, and travel expenses of key agency staff. Fifty-six respondents were interested in corporate data and information, including departmental reports, policies, audits, and evaluations and 49 respondents were interested in data and information about human resources, including descriptions of work, reclassifications of positions, and employment equity.

FIGURE 1.

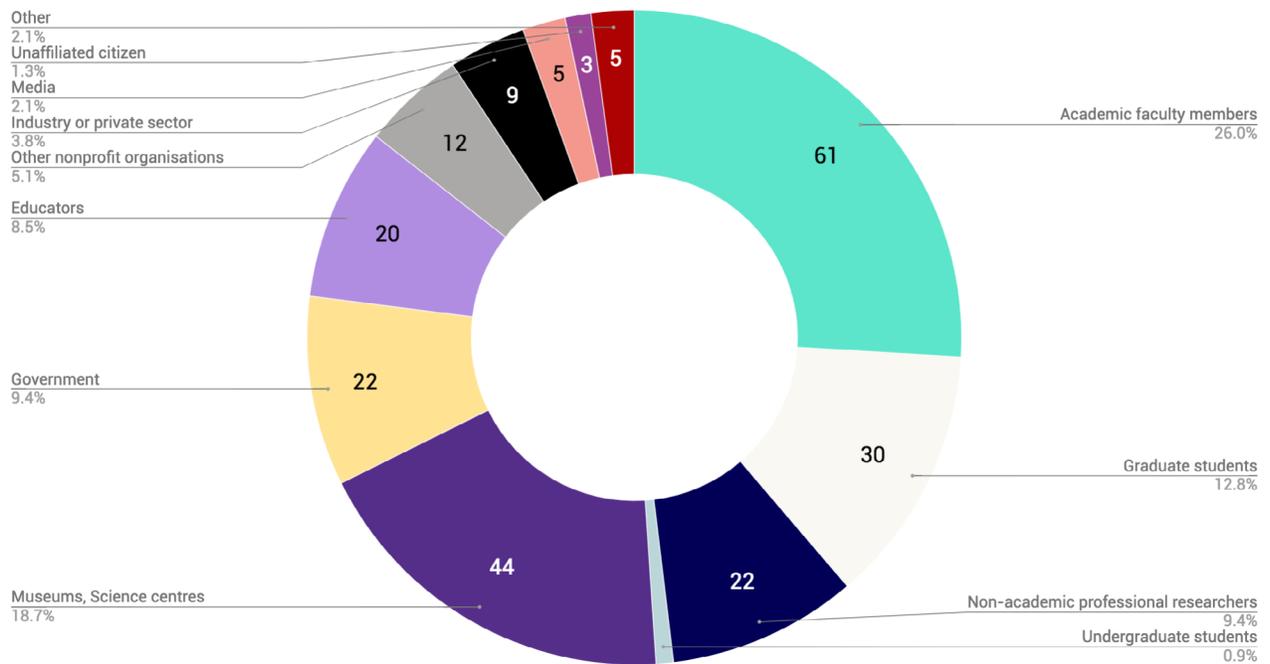
Number of potential users by interest in type of space data and information



Respondents were from a diversity of sectors (Figure 2). The largest representative group were those who self-identified as from the scientific research community, which make up just under half (115) of the total respondents. This group includes academic faculty members (61), non-academic professional researchers (22), graduate students (30), and undergraduate students (2). The next largest group of respondents were from museums and science centres. They comprised of almost 19% with 44 respondents. There were 22 respondents from government, with most of those (19) were from outside of Canada or unknown residency. Twenty respondents self-identified as educators, including teachers and education-focused nonprofits. Twelve respondents were from other nonprofit organisations, nine were from the private sector, five were members of the media, two identified as unaffiliated citizens, and five people described themselves as “other”.

FIGURE 2.

Number of potential users by sector



Potential uses of space data and information

Respondents were asked to describe the potential applications for accessing and using CSA data and information. Below are the different purposes identified by potential users in the questionnaire with specific examples drawn from potential external users.



RESEARCH

Using space data and information for research is a common application for many of the potential user groups, including all levels of the scientific community, those from industry, government, and the nonprofit sector. Research is often not the final objective, but rather a means to another end. The objectives of research may be to publish scholarly articles, but may also be used for the development of new technologies or to help guide policy and decision-making.



DEVELOPMENT OF NEW TECHNOLOGIES

Respondents from the scientific community, the private sector, and the nonprofit sector had an interest in accessing and using space data and information to assist in developing new technologies such as instruments, software, and applications. Data can be used at various stages in the development of these new technologies, from research, prototyping, testing, and training algorithms to using the data directly in their product. Data can also be used to inform business planning and operations.

POTENTIAL USER: Works in private sector (R&D)

INTEREST IN SPACE DATA: To use telemetry data from rovers, rover simulators, and rover prototypes to help test some of the technologies it is developing to improve space systems and missions.





POLICY, PLANNING, AND DECISION-MAKING

Respondents from the scientific community, the private sector, the nonprofit sector, the government, and the media reported a desire to use space data and information in policy, planning, and decision-making. Many potential users wanted to use data for environmental monitoring, evaluation and management. Some examples include disaster management or risk mitigation, resource management, and sea ice mapping.

POTENTIAL USER: Member of the media

INTEREST IN SPACE DATA: To have up-to-date information about the spread of wildfires to help others make decisions about the evacuation of an area or population.



POTENTIAL USER: University-based application developer

INTEREST IN SPACE DATA: To have up-to-date information about the position and thickness of sea ice to help others make decisions about the safety of ice for travel and hunting.



POTENTIAL USER: Government

INTEREST IN SPACE DATA: To use space environment and space weather information to understand how it can affect high-voltage electrical systems to help plan for upgrades to the electrical system to avoid blackouts.





TEACHING AND EDUCATION

Respondents from the scientific community, the nonprofit sector, museums and science centres, and those who self-identified as educators wanted to access and use space data and information for teaching and educational purposes for both children and adults. Age appropriate level data sets were of interest to use for creating interactive demonstrations or performing simple analyses either by the teachers or students.

POTENTIAL USER: Educator (High school)

INTEREST IN SPACE DATA: Access to real data to help graph or chart results of for example, changing temperatures over time to demonstrate the effects of climate change.

In their home province, there are three specific high school units examples of where space data and information might be useful.

- **Grade 9 - Space unit.** Request for Canadian content including what Canada's contributions to the International Space Station are, information on current Canadian astronauts, and current careers related to space sector in Canada.
- **Grade 10 - Weather unit.** Request for data and statistics about weather conditions and changes over time, accompanied by satellite photos to show changes or impact of climate change.
- **Grade 11 - Earth systems and World geography.** Request for data and information on land, water, and air conditions and changes over time.



POTENTIAL USER: Works at a science centre

INTEREST IN SPACE DATA: To use visual and auditory data and information for exhibits, including planetarium theatres or domes.





DISSEMINATION

Respondents from the scientific community, industry and the media were interested in disseminating data, information, and knowledge for wide consumption for the purposes of learning or improving the work of others.

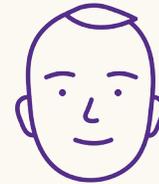
POTENTIAL USER: Member of the media

INTEREST IN SPACE DATA: To tell compelling multimedia stories that would engage readers for a longer period of time using the best possible imagery. For example, satellite data could be used to graphically demonstrate the spread and impact of wildfires.



POTENTIAL USER: Works in private sector

INTEREST IN SPACE DATA: To make Earth observation and atmospheric datasets available on an aggregated data platform product with access that could be free or paid, depending on the source of the data. Openly licensed data are available free of charge, but there is fee for commercially acquired data. The value of this product is in the aggregation of various sources of data to form a more comprehensive product.



OTHER USES

A less frequently mentioned, but nonetheless important potential use of space data and information is for use in hackathons or to help communities through initiatives like Civic Tech or an [Open Data Book Club](#). Events like these provide a space for citizens and professionals to crowdsource solutions for a variety of challenges and societal problems using available data and information. For example, [Civic Tech weekly meetups in Toronto](#) convene people with differing sets of skills to better understand and find solutions to civic challenges through technology, design or other means.

Data and information requirements to facilitate access and use

Data and information requirements across user groups were more similar than different. The largest differences between users would be in their technical capacity to process data and how that affects their needs. High capacity users were more likely to identify as from the scientific community, the private sector, and government. These users have the ability to process raw data with clear documentation to assist, but would prefer pre-processed data to Level 1 or 2. Low capacity users were more likely to identify as educators, from nonprofits, or from the media. These users wanted highly processed data or information that could be easily summarised or transformed into materials for distribution to students, clients, or the general public. Museums and science centres had staff or access to external resources with varying skills sets, but wanted access to processed data for immediate display or data and information that could be easily translated into educational materials.

Below is a table of the common challenges identified by potential users and the data and information requirements that would address these challenges, increase the value, and facilitate use of the data.

USER CHALLENGES	REQUIREMENTS
 LICENSING	
<p>Copyright issues leave users unable to archive, re-publish, and share those data for the use of other scientists who may be able to use the data for another purpose.</p> <p>Licensing language is difficult to understand for the lay user.</p>	<p>Open license preferred</p> <p>Clear licensing language (eg, Creative Commons)</p>
 FREQUENCY OF UPDATES	
<p>Data need to be up-to-date to facilitate citizens or government decision-making, particularly in high-risk situations (eg. ice flows, ice thickness, forest fires).</p> <p>Users want to evaluate changes over time (short and long-term) for resource management.</p>	<p>Most up-to-date and frequently updated (within hours or days)</p> <p>Historical time series data are useful</p>

Data and Information Requirements (Continued)

USER CHALLENGES	REQUIREMENTS
 DATA STANDARDS AND FORMATS	
<p>Data standards and formats vary depending on the field or topic of study. Often international initiatives or partnerships exist which have agreed upon data standards and formats for data publication. Data standardisation increases usability and comparability.</p> <p>Users such as educators or members of the media may not want or have the capacity to process any data products provided.</p>	<p>Aligned with the common practices in each field of scientific study (eg. Earth observation)</p> <p>Common formats for datasets and documents (eg. CSV, JPEG)</p>
 TOOLS AND RESOURCES	
<p>Users need access to additional information to facilitate use of data, particularly users who will process and analyse the data.</p> <p>Tools used to view or analyse data are expensive and prohibitive.</p>	<p>Accompanying materials to the data, such as data dictionaries, metadata, calibration and data quality documentation.</p> <p>Preference for open source software</p>
 DISCOVERABILITY AND ACCESSIBILITY	
<p>Most potential users were unaware of Canada’s Open Government Portal nor were they familiar with resources available on the CSA website. In each instance, data that were currently being used by interviewees, were not accessed through CSA.</p> <p>Inconsistent or slow internet connectivity in remote regions or developing countries as a barrier to being able to download high resolution products and large subsets of datasets.</p>	<p>Easy to find - High quality metadata are necessary; a centralised portal; and targeted materials by age or sector</p> <p>Easy to download - Zipped and small batches are made available; options for high and low resolution products. useful</p>

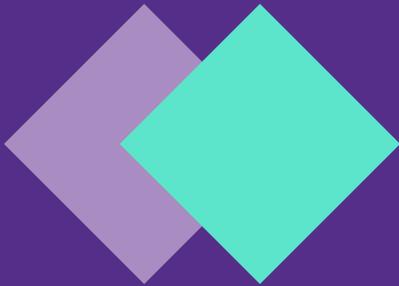
Providing multiple products for the same asset, building tools to view and analyse data, or clearing the rights over data not owned or held by CSA will be dependant on the resources available. Establishing the priorities of CSA to provide these additional services would be facilitated by an articulated data management policy. CSA' s Open Government team is currently lacking clarity on where to focus its efforts once it has complied with its Open Government obligations. For example, should it focus on increasing accessibility and engagement with CSA's open data assets, or should it focus on sharing more science data by clearing the rights of third-parties, or providing access data via third-party websites?

Without accompanying documentation, including metadata, the value of CSA's data and information assets is severely compromised. Improving discoverability and accessibility of its assets, even those that are not yet published should be a priority of CSA. Potential users can discover datasets, evaluate if it will meet their needs, and request the data if needed. In addition, when resources permit, small batches of data and low resolution products should be provided alongside larger or complete batches of data and high resolution products.

To further increase discoverability of data, CSA should engage more directly with its potential data and information users through collaboration with its communications team and other teams at CSA. For example, using social media, news releases, and blog posts to inform potential data users about recently released assets or organising a hackathon where participants will work with real CSA data. The consultation described in this chapter was a method for CSA to engage with potential data and information users, and while initiated by the Open Government team, it involved collaboration with CSA policy shops, communications, and the three program sectors.

CSA's central challenge to prioritising its data and information sharing activities is a policy gap between what CSA is mandated to do via its Open Government obligations and how it would like to better accomplish the agency's broader mission by opening more CSA-funded science data. Lacking an agency-wide data management policy, the CSA Open Government team is unable to reconcile its competing priorities.

The first section focused on evaluating CSA's current data sharing practices compared to data sharing practices with other space agencies. This section presented results of a consultation with potential space data and information users to better understand their needs. The next part will conclude this report and provide three scenarios aimed at addressing this policy gap identified above.



PART 3

CONCLUSIONS AND RECOMMENDATIONS



RECOMMENDATIONS

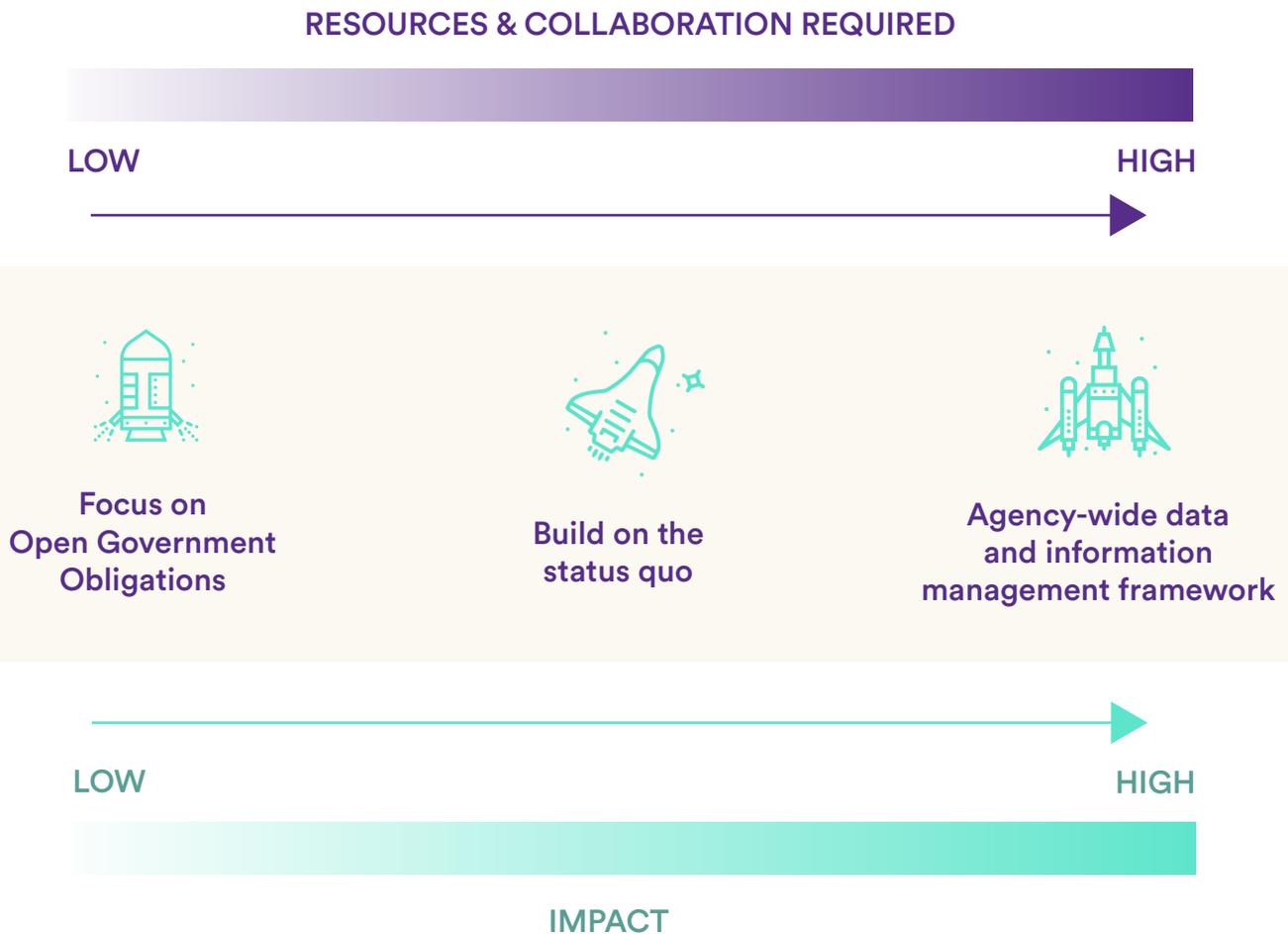
Powered by Data was hired by the Open Government team at CSA to give advice and guidance to improve the ways it can share more data and information to further CSA's mission. CSA has been sharing science data and information that it holds through the Open Government Portal. To date, 16 scientific datasets have been shared on the portal. This practice has been limited because CSA does not own much of the data that are produced from projects that it funds. Data that are produced with CSA funding through grants and contributions are usually owned by the third-party. In addition, data management plans are not a standard practice of these agreements. To share additional scientific data that CSA has contributed to and have been made available online on third-party websites, the CSA Open Government team has initiated a space on the CSA website to provide the links to these external sites.

The CSA Open Government obligations are limited to opening data and information assets which it owns. This work is done often with already limited resources. However, CSA recognises that open data and information is one way to give new life to science data and maximise the return on CSA investments made through grants and contributions and contracts, and enable innovation, economic development, collaboration, and evidence-based decision-making in the sector.

Open data is one tool that can help CSA better accomplish its mission to advance the knowledge of space through science and to ensure that space science and technology provide social and economic benefits for Canadians. However, disseminating additional resources beyond the Open Government Portal is not a mandated activity for the Open Government team at CSA. This activity is in competition with providing additional digital services, such as tools or applications aimed to increase access and use of CSA's open data and information assets. CSA is lacking a data management policy that would help it establish data sharing priorities, create processes for, encourage collaboration on, allocate resources to, and make decisions about the collection, management, and archiving of data and information.

Within the context that more open scientific data and information at CSA would help fulfil CSA's Open Government obligations and its mission, but that the CSA Open Government

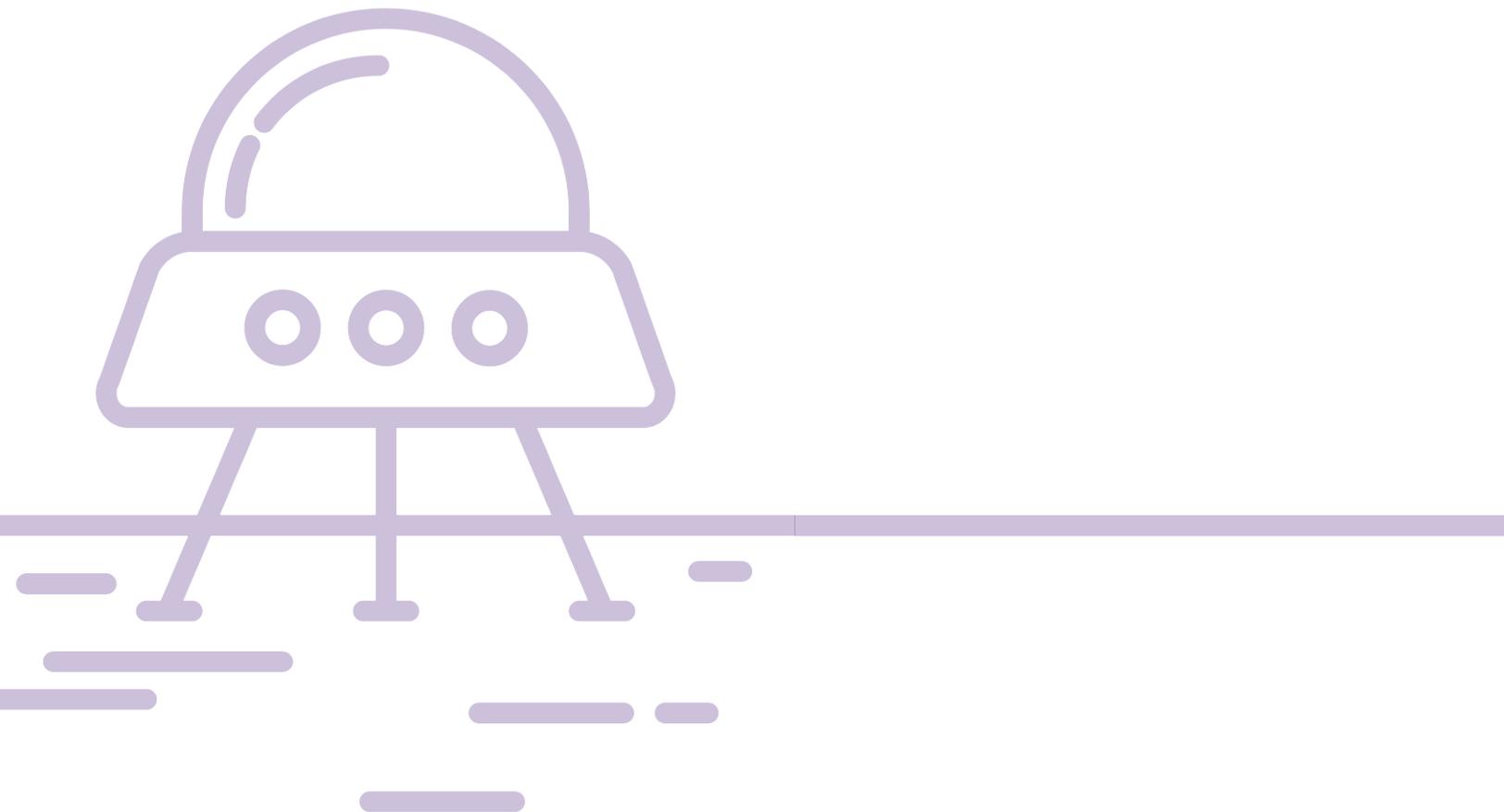
team does not have all the tools available to it to open data that CSA does not own, we propose three scenarios that require an increasing amount of collaboration with internal and external partners, and an increasing amount of resources. However, the more collaboration and resources required, the higher the level of potential impact.



The first option, “focus on Open Government obligations” has the Open Government team using open government to accomplish its mission by concentrating its efforts on data and information CSA owns. The second option, “build on the status quo”, involves the CSA Open Government team continuing and improving the work it is currently doing: knowledge dissemination for both CSA owned assets, but also those scientific data and information to which it contributes, but does not own. The third option, “Agency-wide data and information management framework” involves CSA providing a framework for data and information sharing that would provide the mandate for any change to policies and practice throughout the organisation.

These three scenarios are not distinct options but exist on a scale of possibilities that could be adopted in a non consecutive order. Furthermore, the recommendations are not mutually exclusive to one scenario, but could be implemented cumulatively. For example, CSA could focus on its Open Government obligations and implement the recommendations listed under that scenario and then begin developing an agency-wide data management plan next year and continue to implement recommendations from all three scenarios. Alternatively, CSA could begin to develop an agency-wide policy now and implement the recommendations of scenario 1 and 2 in the meantime.

Each scenario is presented below with advantages and disadvantages to help CSA consider the consequences of each. Recommendations for each scenario are intended to help maximise the impact of each option. The report will conclude with some advice on how to decide next steps.



◆ LOW LEVEL OF COLLABORATION & RESOURCES, LOW IMPACT

1. Focus on Open Government obligations



Get to where you need to go without the bells and whistles.

Focus on data and information that CSA owns when updating its inventory and releasing data on the Open Government Portal. This was the original mandate of the CSA Open Government team in 2016.

Pros

- Minimal risk of noncompliance to Canada's Open Government policy.
- Requires minimal investment.
- Requires relatively little collaboration with other teams across CSA.
- Potentially helps CSA achieve the Open Government goal of increasing trust in government through transparency.

Cons

- Missed opportunity to use open data & open science to promote CSA's domestic and international contributions to the space sector.
- Because CSA does not own the majority of data and information generated through its contributions, CSA will not receive most of the benefits that the Open Government policy aims to achieve by only complying with the letter of the policy.
- Does not provide CSA the capacity or mandate to expand open science practices.
- Unlikely to result in CSA increasing its capacity to leverage data to better accomplish its mission.

Recommendations

- Assets published on the Open Government Portal should only be data and information that CSA owns or over which it has the rights.
- Improve the open data inventory by publishing additional optional fields and ensuring frequent updates.
- Improve the quality of open data by ensuring the publication of appropriate supporting documentation, such as data dictionaries, metadata, calibration and data quality documentation alongside the dataset.

◆ MEDIUM LEVEL OF COLLABORATION & RESOURCES, MID-IMPACT

2. Build on the status quo



Shoot for the stars, despite modest resources.

CSA continues and expands its practice of providing access to space data and information assets beyond that of which it owns including open space data and information from partnerships with third parties. The CSA Open Government team has been pursuing these types of activities without a confirmed mandate or policy guidance.

Pros

- Consistent with the spirit of Canada's Open Government obligations.
- Through increased activity, CSA is more likely to accomplish the goals of Canada's Open obligations, including increased innovation, collaboration, and economic development.
- Help to foster learning and a culture of open data and open science across CSA and develop its capacity to leverage data to better accomplish its mission.
- Promotes CSA's contribution to the space sector by leveraging the most impactful data and information assets, not just those which are owned by CSA.
- Highlights the work of CSA's partners and thus contributes to promoting open science.

Cons

- Current level of investment only allows for modest success.
- A lack of clear policy guidance makes it difficult to measure success.
- Lack of consistency: Other parts of the organisation will continue to make different decisions about when and to what extent data are made available.
- "Coalition of the willing": In the absence of a clear, organisation-wide data management policy, collaboration with or support from other units at CSA will be managed on a case-by-case basis.

Recommendations

- In the absence of clear policy guidance, a committee should be created and led by the Open Government team. It should include participation from project managers from each of the three program sectors. The main goals of the committee would be to share information about data management activities and initiatives, specifically the scientific data being produced in each program sector and any data mobilisation activities. The committee would also promote open

data practices among the project managers.

- Prioritise high-impact data and information based on CSA mission and user feedback.
- Publish data and information, including links to partnership data archives in a centralised place on the CSA website.
- Improve discoverability of data by publishing extensive metadata, including for data and information assets that cannot be made openly available due to various restrictions.
- Continue efforts to clear rights over additional high-impact datasets.
- Increase efforts to provide different versions of the same dataset to deepen impact and meet the needs of multiple users.



3. Agency-wide data and information management framework



A framework for collaboration across the organisation to reach new heights.

With an articulated top-down agency-wide data and information management framework, CSA staff will have clear directions on what principles and organisational goals should drive its priorities, policies, and practices regarding data collection, sharing, and use.

Pros

- CSA will be able to share more federally-funded science data and information and contribute to achieving the federal government's open science objectives.
- Likely to accomplish both the Open Government and the Open Science goals, including, increased innovation, collaboration, and economic development in the space sector.
- Enables CSA to use its data to contribute more actively in the domestic and international space sectors.
- This highly collaborative approach will result in more effective and consistent data management practices across CSA that will better leverage data to accomplish the agency's mission.

Cons

- Not cost neutral: Implementing would most likely require both one-time and ongoing investments in infrastructure and infrastructure management. This is regardless of whether the work is done internally or through partnerships.
- Organisational change requires ongoing support and engagement from senior management.

Recommendations

- Develop and implement a horizontal framework that lays out the vision, objectives, and principles of CSA's approach to data management.

- Establish joint leadership to develop and implement this framework between CSA's CIO and its upcoming Science Advisor.
- Adopt an open science approach and establish guidelines for reasonable exclusions that align with the responsibilities of the agency and its mission. Eg. data embargos, commercialisation opportunities, confidentiality.
- Data management plans (DMPs) should be a standard in project management, grants and contribution agreements, contracts, and CSA's own missions.
- Provide workshops and trainings across all levels of the organisation to increase open data knowledge and capacity.



CONSIDERATIONS AND NEXT STEPS

The two largest considerations about the feasibility of each scenario are the potential for organisational support and the related availability of resources.

CSA is expected to comply with the federal Open Government policy, which aims to open data and information that are owned or held at CSA with priority for assets of high business value. As a science-based agency, data and information resulting from CSA funding are considered high priority, but are not frequently owned or held by CSA due to the current standard practices for grants and contribution agreements and contracts.

It is clear that there is some appetite to expand CSA's practice of opening CSA-funded science data and information, as evidenced by the:

- Participation of CSA program sector staff in various open data sharing initiatives (eg. GEOSS, CODATA, Open Government working groups, Open Data Day at CSA);
- [CSA's Go Canada open by default data policy](#);
- Attempts initiated by CSA or third-parties to clear the rights over data and archive them for CSA to release openly (eg. MOST, WINDII);
- Efforts to continue to scan and release Alouette data; and
- CSA's 2007 Science data workshop organised by program sector staff to raise the level of discussion about space science data in Canada and to encourage the development of a science data community that would cut across the traditional space science disciplines.

In addition, the CSA Open Government team, in collaboration with program sector staff and communications has begun to share CSA-funded science data hosted on third-party websites via CSA's website. The website already had some data and information with more restrictive licenses available.

The potential for organisational support is strong for the first scenario because CSA must comply with these obligations. The status quo has been receiving some organisational support in the form of collaboration between the Open

Government team and some program sectors, but without a clear data management plan expanding the practice will be decided on a case-by-case basis with little consistency. An agency-wide data management plan requires organisational support, particularly from the top, to provide clarity on the vision and principles for the fundamental questions around data. Senior management also plays a key role in change management and adopting an organisational-wide culture around data.

Similarly, the scenarios above each require a different level of investment in terms of finances and human resources. Human resources include more time and more time of more people across the organisation. To focus on only Open Government obligations requires minimal collaboration with program staff and communications, while building on the status quo or a agency-wide data management framework each require increased involvement with more internal and external stakeholders. The scenarios are also not cost neutral and each scenario costs increasingly more, with investments in infrastructure and infrastructure management.

Because of the evolution of government priorities in regards to open data and open science, the next step for CSA's Open Government team is to seek clarification about CSA's organisational priorities in relation to data and information management, with a particular focus on the role of open science in helping achieve CSA's mission. The role of open data in enabling innovation and social and economic development is growing quickly. To more efficiently leverage its data to achieve these benefits, the CSA Open Government team needs clarification on what the scale of its role in the space and science sector will be. CSA will continue to open science data and information in which it owns as part of its Open Government obligations, but requires clarification as to what degree will it increase its efforts to open more data and information from CSA-funded science activities.