Columbia Basin Water Hub

Data Management Plan

and Operational Framework

Living Lakes Canada

Created By: Santiago Botero and Paige Thurston
Date Created: March 2021

Updated By: Santiago Botero and Paige Thurston
Date of Last Update: August 2021
Version: 1.2

Executive Summary

Over the past two decades, a series of reports have projected the impacts of climate change on hydrological cycles in the Canadian Columbia Basin, which are occuring.

A 2017 conference and follow up research focused on Columbia Basin water data, concluded there were significant water data gaps and therefore insufficient knowledge of the Basin's watersheds to make informed water management decisions to increase climate adaptation options. There was consensus on the need for coordinated monitoring and for a central repository to store and access current and historical water data collected.

The Water Hub aims to bring together stewardship groups, First Nations, government, industry, and academia in the Columbia Basin to coordinate a collaborative approach to water data collection, analysis, sharing and storage for the future. At the Water Hub, we believe that the power of big data, data analytics, cross-sector partnerships, and openness are the steps towards ensuring that we will have water resources, and their associated ecosystems, for future generations.

The Water Hub is just one element of a larger initiative to create a structured scientific approach to water monitoring within the Columbia Basin. This will be achieved through the Columbia Basin Water Monitoring Collaborative, which will eventually be its own legal entity with a board of trustees. The Collaborative aims to provide funding, research opportunities, and training, as well as communicating information to the public and government.

The Water Hub's development was guided by steering committees that included all levels of government, First Nations, community-based monitoring groups, professional consultants and the public.

Data will be obtained from diverse sources, including community-based water monitoring groups and local governments, and additional links will be provided to data on other databases. The Water Hub will host a variety of types of data. Tabular files (csv, xlsx), PDFs, images, maps, geodatabases, and any other file format are able to be shared. CKAN displays data in the form of Datasets, which may contain one or more Resources (files). Resources may be in different formats, for example a CSV spreadsheet of data, a JPEG site photo, and a PDF report which interprets the data. On the site, metadata will be displayed at the Dataset level, and more specific metadata will be displayed at the Resource level.

The data sharing process has been designed to be as simple as possible. In the early phases, the upload process will be carried out by Water Hub staff. In the future, once a training program has been developed, contributors may begin to upload their own data. Templates have been developed to facilitate this process and promote consistency.

Each contributing organization will complete a Data Sharing Agreement with the Water Hub and a Data Management Plan. Contributors maintain ownership of their data, and if certain criteria are met, data can be designated as "Private" and visible only to members of that organization.

All data will undergo a QA/QC process. With consultation from the community, specific QA/QC protocols have been developed for different types of data. Additionally, all data will be assigned a Water Hub Data Grade, which informs users of the quality of the data. This will ensure that a resource is of appropriate quality for the user's purpose.

Through DataCite, a Digital Object Identifier (DOI) will be minted for each dataset. DOI's will allow the Water Hub to track the citation of data. This will provide evidence of data being used for decision making or research.

The Water Hub is a flexible program which aims to adapt over time to meet the needs of the water stewardship community.

Acronyms and Abbreviations

CABIN	Canadian Aquatic Biomonitoring Network
CBT	Columbia Basin Trust
CKAN	Comprehensive Knowledge Archive Network
FAIR	Findable, Accessible, Interoperable, and Reusable
ISO	International Organization for Standardization
LLC	Living Lakes Canada
NAS	Network Attached Storage
RISC	Resources Information Standards Committee
QA/QC	Quality Assurance and Quality Control
WSC	Water Survey of Canada

Table of Contents

1.0 Introduction

The Columbia Basin Water Hub (Water Hub) is a central repository for water data collected within the Upper (Canadian) Columbia Basin. For our purposes, “water data” includes any data regarding water that is collected from: streams, lakes, snow, glaciers, groundwater, wetlands, and riparian zones. This could be in the form of tabular, numerical data; descriptions of cultural values; images; reports or other formats. The Water Hub is the confluence of all the water data that is increasingly needed for science-based decision making and in the creation of climate resilient communities. By adopting open data principles, and maintaining a standard of data quality, the Water Hub is a place where anyone can access accurate data regarding their watersheds, increasing water literacy within our communities.

Motivation

The Water Hub aims to bring together stewardship groups, First Nations, government, industry, and academia in the Columbia Basin to coordinate a collaborative approach to water data collection, analysis, sharing and storage for the future. At the Water Hub, we believe that the power of big data, data analytics, cross-sector partnerships, and openness are the steps towards ensuring that we will have water resources, and their associated ecosystems, for future generations.

Context

Publications relevant to climate change within the Columbia Basin are listed below. These reports raise concerns regarding changes in the hydrological cycle for this region, which will have various environmental, economic and social impacts on the Columbia Basin.

Research by Dr. Martin Carter for the Columbia Basin Trust (CBT) shows that data gaps exist within the Basin (2017). Another paper by Carver explains the need to coordinate water monitoring in the Basin (2019).

Guiding Reports

- Pacific Climate Impacts Consortium. 2006. Preliminary Analysis of Climate Variability and Change in the Canadian Columbia River Basin: Focus on Water Resources. Victoria, BC: University of Victoria.
- Pacific Climate Impacts Consortium. 2013. Climate Extremes in the Canadian Columbia Basin: A Preliminary Assessment. Victoria, BC: University of Victoria.
- Carver M. 2017. Water Monitoring and Climate in the Upper Columbia Basin, Summary of Current Status and Opportunities. Prepared for Columbia Basin Trust.
- Carver M. 2019. Water Monitoring and Climate Change in the Upper Columbia Basin-Guidance Information for Planning Monitoring Programs. Prepared for Columbia Basin Trust

In light of these reports, in 2017 Living Lakes Canada co-hosted a conference called *Cracking the Code in 3D: Water Data Hub and Monitoring Framework* with First Nations, government, and experts from around North America to discuss a possible approach to water stewardship in the basin. That is where the Water Hub and Columbia Basin Water Monitoring Collaborative (Collaborative) concepts originated.

In June 2020, Living Lakes held a workshop called *Expanding Water Monitoring Within Canada's Upper Columbia Basin*, led by Carver and Greg Utzig. Various senior hydrologists discussed the best possible scientific approach to water monitoring within the Columbia Basin. The proceedings from this workshop can be found through the link below.
https://livinglakescanada.ca/wp-content/uploads/2020/10/LLC-wm-wkshp_prcdngs-final_Oct-08-2020.pdf

The Columbia Basin Water Hub was created in response to these reports and meetings, with the goal of facilitating a collaborative approach to water data collection, analysis, and sharing.

The Water Hub is just one element of a larger initiative to create a structured scientific approach to water monitoring within the Columbia Basin (Figure 1). This will be achieved through the Columbia Basin Water Monitoring Collaborative, which will eventually be its own legal entity with a board of trustees. The Collaborative aims to provide funding, research opportunities, and training, as well as communicating information to the public and government.

The aim of the Water Hub is to provide open data to empower a water literate constituency. This means providing the public with information that is easily digestible and useful.

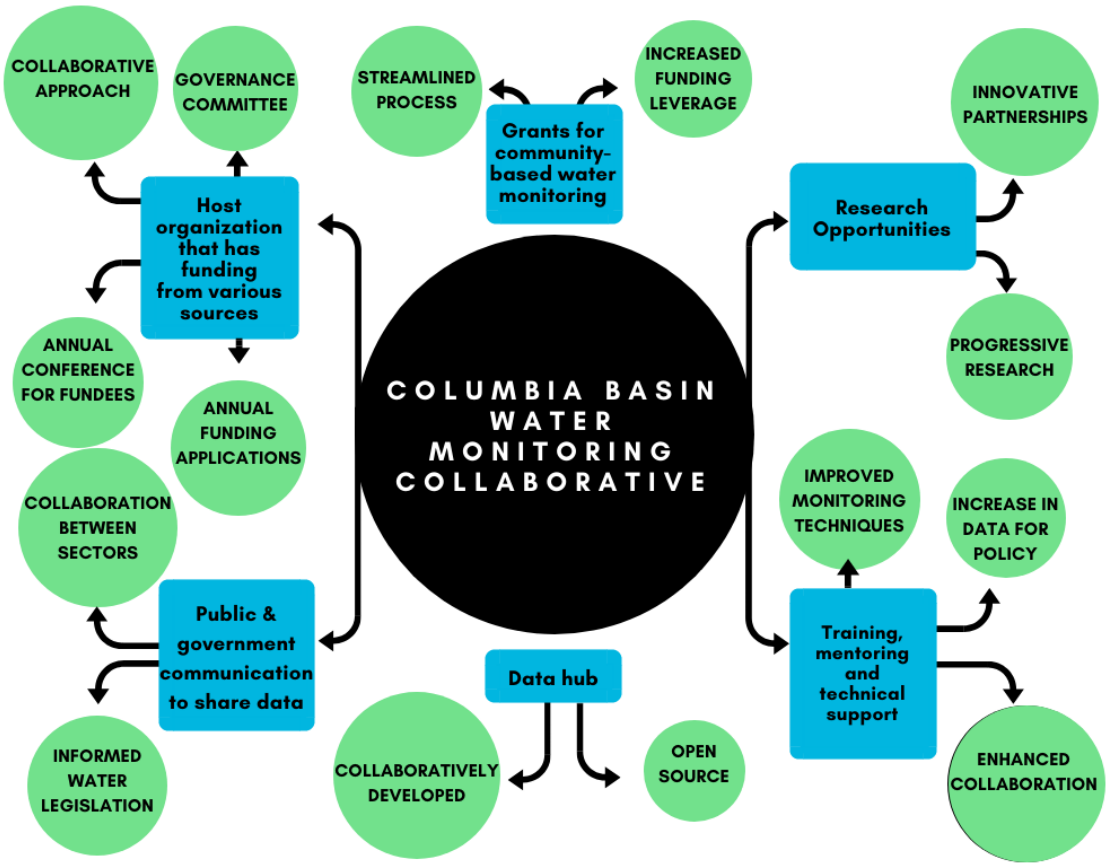

Figure 1: Components of an initiative to create a structured scientific approach to water monitoring within the Columbia Basin, with the Water Hub as one element of this.

External coordination

Right now, the external coordination is provided through guidance from LLC. In the future, it would be beneficial to create a dedicated board for “technology guidance” which may give guidance in regards to where technology advances are going, and provide guidance for further development.

Governance

Governance will eventually fall under the governance structure of the Columbia Basin Water Monitoring Collaborative. Guided by a science board, and a board of trustees, the Water Hub will be guided by the needs identified by the Collaborative.

2.0 Database and Development

How the Database was Built

The Water Hub runs using CKAN (Comprehensive Knowledge Archive Network). CKAN is an open source database service which uses instances (or store fronts) to display data which is held in CKAN's DataStore. The Water Hub is a customized CKAN installation (https://ckan.org/). CKAN was installed by following the documentation at this page. The version running on the server is 2.8.2. Some pages in the documentation references version 2.7.x, but it is largely the same as 2.8.x.

Due to security issues with the officially supported version of Solr, a newer version of Solr (6.5.x) was installed and some modifications were made for compatibility with CKAN, as described here. It is not necessary to install Solr locally, it can simply be run from the source folder.

Our CKAN instance has a variety of custom extensions, with the most important ones being the custom schema, mapsearch, and xlsload. More information can be found in the developer notes (only accessible through Living Lakes network, if you would like to see the developer notes please email current Database Manager santiago@livinglakescanada.ca).

Developer Contact Information

The Water Hub CKAN site was developed by Genki Maps, a geospatial software consultancy:

Michael Peterman

Web: www.genkimaps.com

Email: genkimaps@gmail.com

Phone: 778-806-0360

Living Lakes Web Admin

Website Admin for Living Lakes:

Shawn Wernig

Email: shawn@eggplantstudios.ca

Phone: 250-688-8008

Shawn has been given sudo access to the Water Hub server via SSH. Shawn has access to the Living Lakes domain host, and was responsible for pointing the subdomain waterhub.livinglakescanada.ca to the server hosting the Water Hub, and adding the DNS records to support Mailgun email notifications.

Server Hosting

The Water Hub runs on a server hosted by Carl Data solutions, the makers of FlowWorks. Living Lakes has an agreement with Carl Data solutions to provide this server.

Contact at Carl Data Solutions:

Mike McDonald, VP of Engineering

Email: admin@carlsolutions.com

www.carlsolutions.com

Data Storage

Data is currently being stored in Microsoft servers in Montreal, and our CKAN instance is running on a virtual machine hosted by Carl Data. The way that CKAN works is through the DataStore, which is where the data is stored, and then the metadata is shown through our site/instance. An instance is a version of CKAN which has been personalized for various uses.

3.0 The Water Hub Site

The Columbia Basin Water Hub is bridging the gap between data collection and a long term access and storage solution. Currently, there is important data that needs to be housed in a centralized location. A diverse range of organizations carry out a variety of water monitoring programs in the Basin. A brief summary of both current and previous water monitoring efforts in the Columbia Basin is summarized in *Water Monitoring and Climate in the Upper Columbia Basin, Summary of Current Status and Opportunities* (Carver, 2017). It is through this diversity of data collection that the Water Hub will have an influx of data. Data uploads will take the form of the database staff uploading datasets provided by monitoring organizations and institutions, and in the future, users trained under the "Water Hub Certification" program that will be developed. Certified users will be able to carry out their own QA/QC and upload data to the Water Hub.

The Water Hub adheres to FAIR (Findable, Accessible, Interoperable, and Reusable) Data Principles, which are the backbone of open data.

Types of Data Hosted

The Water Hub will host a variety of types of data. Tabular files (csv, xlsx), PDFs, images, maps, geodatabases, and any other file format are able to be shared. Here is where CKAN's flexibility helps, allowing the Water Hub to host all of the different shapes that water data takes. CKAN displays data in the form of Datasets, which may contain one or more Resources (files). Resources may be in different formats, for example a CSV spreadsheet of data, and a PDF report which interprets the data.

Data Sources

Data collection is conducted by a variety of different organizations. There are many community stewardship groups which have water monitoring programs in place. Additionally, data is collected by different levels of government, and community water systems. These community-based water monitoring programs are locally focused, and reflect the specific concerns from the community. Through existing connections, and through ongoing outreach, the Water Hub team will connect with these monitoring groups and invite them to share their data through the Water Hub.

Another source of data will arise through the partnerships created with industry partners. Ideally, the Water Hub will create relationships with industry leaders in forestry, mining, and agriculture.

The Water Hub also wants to offer Indigenous Nations the opportunity to have a platform where, if they so choose, they are able to display knowledge in ways which are specific to each community.

Data which is already publicly available through other online sources (such as Provincial databases) will be incorporated into the Water Hub in the form of links to the existing external databases.

Metadata

Metadata will be displayed for each dataset, and more specific metadata will be displayed for each resource within the dataset. The current metadata schema includes:

Dataset Level Metadata

- Dataset Name
- Abstract
- Data Upload Organization
- Visibility
- Data Collection Organization
- Data Steward Email
- Citation
- Dataset Upload Year
- Funding Description
- Data Disclaimer
- Other Data Sources
- Brief Description of Location
- DOI
- Start Date
- End Date
- Licensing and Attribution
- Data Type (type of waterbody)
- Keywords
- Latitude
- Longitude

Resource Level Metadata

- File Upload or Link
- Header Row
- Resource Name
- Citation
- Format
- Description of Data
- Resource Location
- Data Collection Information
- Water Hub Data Grade
- Data Processing
- Data Disclaimer
- Water Hub Certified Upload

Organization of Data on Site

CKAN offers a number of tools to organize, sort and view resources and datasets, including Groups and Keywords.

The Water Hub uses the Groups function to categorize datasets by the Basin Hydrologic Regions identified in the 2019 report by Dr. Martin Carver.

The Map Search displays Datasets categorized by Data Type, the type of waterbody that the data is related to.

For Numerical data, which may include Spatial files with Numerical data, the Water Hub team uses a set of parameters developed with input from the community to assign a Water Hub Data Grade to each resource (see QA/QC Procedures). Other types of resources, such as written reports, Indigenous knowledge, and maps are not classified using these parameters. This grade is included in the Data Grade field of the Resource metadata, and also included as a keyword for the Dataset (Datasets which include Resources of different grades will display all applicable Data Grades as keywords).

Multiple Keywords can be applied to a dataset, and can be used for searching or filtering searches. Keywords include the Water Hub Data Grade for each resource contained in a dataset, and can also include attributes like Water Quality or Community-Based monitoring.

Each organization that shares or uploads data to the Water Hub also has the opportunity to create their own Organization Page, which includes Dataset, About and Activity Stream sections. This function can be used to view all data that has been shared by that Organization. Groups contributing data who do not wish to maintain their own Organization Page, or who are no longer active can have their data uploaded using the Columbia Basin Water Hub Organization, and the Water Hub will act as the steward of the data.

DOI Minting

Another feature of the Water Hub is that DOIs will be assigned to each dataset that is hosted on the Water Hub. To implement this, LLC has become a member of the DataCite Canada Consortium, a partnership between the Canadian Research Knowledge Network (CRKN) and Canadian Association of Research Libraries (CARL) Portage Network, which is responsible for minting DOIs in Canada. DataCite Canada is the Canadian chapter of DataCite which is a global network of DOI minters.

“DOIs are persistent identifiers for digital objects which are widely used for identifying published content, datasets, and other scholarly research outputs,” (CRKN, 2021). DOI’s will allow the Water Hub to track the citation of data. This will provide evidence of data being used for decision making or research.

In this beginning phase, DOIs will be generated manually by Water Hub staff using the DataCite Fabrica interface. In the future it is possible that this process could be automated using a CKAN extension. An example can be found here: https://github.com/NaturalHistoryMuseum/ckanext-doi

Once a dataset has been uploaded to the Water Hub, the DOI can be minted, and will be displayed in the Dataset Metadata on the site. This means that the dataset is also findable through Fabrica, which will increase the access and usage of the data.

4.0 The Data Management Framework

Data Management Plans

With the guidance of the Water Hub team, each data contributor will complete a Data Management Plan using templates developed by the Water Hub team. This consists of two sections. The first, Organization Overview, will describe the organization's water monitoring programs, and data management practices, and clarify the expectations of both parties with regard to data sharing and collaboration. This section will be completed at the beginning of each organization's involvement with the Water Hub, and will be updated over time to reflect the most current information.

The second section, Monitoring Program Details, involves more detailed information about a water monitoring program or dataset, including objectives, QA/QC procedures, methods and instruments used. This will be completed for each dataset or monitoring program for which an organization plans to share data. Data from a single monitoring program may be displayed using separate datasets for each monitoring location, but only one copy of this section needs to be created if the parameters and methods are similar across the entire program. If an organization is sharing data from different monitoring programs, or if the parameters and methods vary considerably, they will complete a separate copy of the Monitoring Program Details for each program. Groups who are uploading their own data only complete the Organization Overview portion of the Data Management Plan, as they will enter the details of the monitoring program directly into the upload template themselves.

As these plans describe the organization's monitoring programs, instruments, methods and data management practices, potential issues with data quality may be discovered at this stage. Options to improve the quality of data collection or management can be discussed with the organization at this time.

Data Management Flow

The Water Hub shares data in two different forms (Figure 2). First is data that is hosted by the Water Hub. This data may be sourced from local monitoring groups or citizen science initiatives. These datasets are either provided to the Water Hub in their raw format, or they are uploaded by groups and volunteers that have achieved the "Water Hub Certification".

The second stream of data consists of links to resources available on other databases or websites that are systematically searched for and added by the Water Hub team.

Hosted Data

Data that is hosted by the Water Hub will initially be submitted directly to Water Hub staff who will conduct QA/QC and upload the data. Once a training program has been developed, monitoring groups may begin to upload their own data, with the Water Hub team conducting QA/QC and assigning a grade after data is uploaded.

Once the Water Hub receives a dataset, it is inventoried in the Data Tracking Spreadsheet and stored in the Incoming Data folder in the main Data Management Folder, within a subfolder that is also accessible to the submitting organization (organizations may upload their data directly to this subfolder). When a Water Hub employee begins to process the file, a copy is saved to the Working Data folder and the spreadsheet is updated. At this stage, Water Hub staff carry out the QA/QC protocols, fix any issues and format the file for upload. Once the dataset is ready for upload, it will be moved to

the Finished Data folder and the inventory spreadsheet gets updated. Files in the Finished Data folder will be uploaded to the Water Hub by staff, and also backed up on the Living Lakes NAS drives (please see Back Up/ Archiving Section).

Templates have been developed to promote consistent formatting, inclusion of all relevant metadata and to streamline the upload process. These can be used by Water Hub staff to format files provided by organizations, or the templates can be provided to organizations to complete themselves once appropriate training has been completed.

Linked Data

Relevant data that is already available on another database will appear in the Water Hub in the forms of links. This data will also undergo a QA/QC process developed for linked data. Once linked datasets have undergone the QA/QC process, they will be linked directly to the Water Hub.

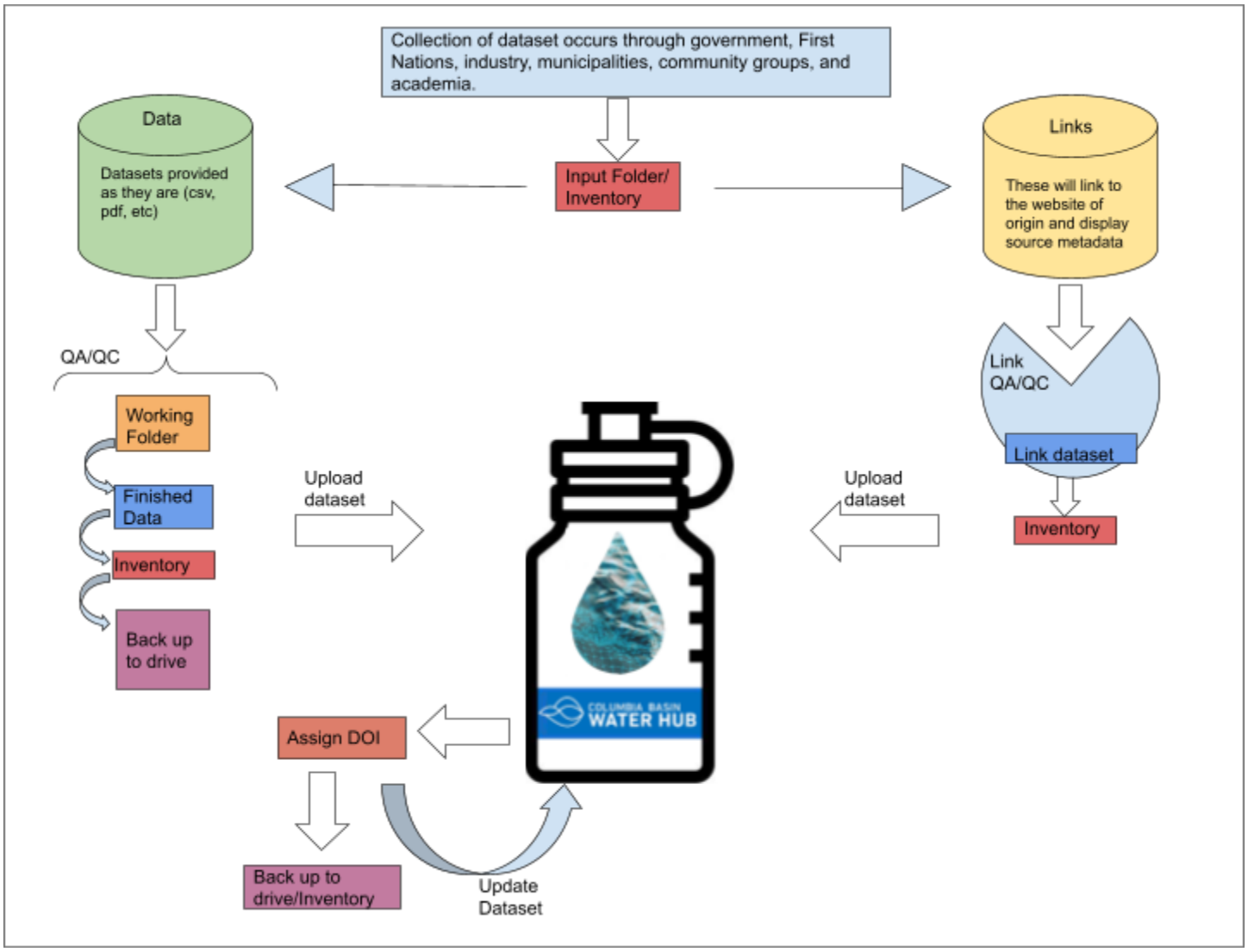

Figure 2: The flow of data from collection to becoming available in the Water Hub.

QA/QC Procedures

Procedures

All data which is uploaded or linked to the Water Hub will undergo some form of QA/QC protocol, but the process will differ depending on the type of data. Five QA/QC Procedures have been developed:

1. Procedure for numerical data hosted and uploaded by the Water Hub

2. Procedure for data collected using an existing protocol, hosted and uploaded by the Water Hub

3. Procedure for reports and other non-numerical forms of information

4. Procedure for data uploaded by users, with QA/QC by the Water Hub

5. Procedure for data linked to other databases

The steps for each QA/QC procedure are detailed in the Water Hub QA/QC and Grade Procedures document.

Metadata completeness

A primary focus of the Water Hub QA/QC system is the provision of complete metadata schemas at the Dataset, Resource and Observation level.

Water Hub staff work with contributors to ensure that all metadata fields in the upload templates are completed. The use of standardized templates (different templates are available for different types of measurement) will ensure that metadata is displayed in a consistent manner with respect to format and units. The templates utilize data validation for certain fields to promote consistency.

In addition to metadata being displayed on the Water Hib, the metadata fields are also included in the uploaded files wherever possible. Including the dataset and resource metadata in the files themselves will also ensure that the metadata stays with the data when it is downloaded, to decrease the likelihood of this information being lost or mistranscribed by a user.

To eliminate errors sourced from transcription or transfer of data, it is suggested that organizations use these templates to collect and store their data, but this is not mandatory as many organizations will already have their own preferred tools.

Water Hub Data Grade Classification

Using the standards developed through consultation with the community and adapted from existing Provincial standards, Water Hub staff will assess each resource to determine its Grade. Table 5 contains the definitions of each Data Grade and the type of data that they are associated with. Figure 3 illustrates the decision making process used to assign a Grade to a resource.

Much of the data being collected in the Basin is of a high standard, but may not conform to an existing standard such as RISC. The intention of the “Technical Research Grade” designation is to communicate to users which resources have achieved a level of quality that would make them suitable for purposes such as scientific use or decision making.

While datasets which do not meet the criteria to be designated as Technical Research Grade may still contain very useful information, they should be used with the caveat that important information may be missing, or the accuracy of the data may be in question. This system of classification will ensure that the quality of data being used reflects what is needed for the circumstances.

Table 5: Water Hub Data Grade definitions and the type of data that they are associated with.

Grade	Data Type
CABIN/ISO/RISC/WSC: Data which has been collected according to an existing standard method will simply display the applicable standard.	For numerical resources, and spatial resources (geodatabases, shapefiles,etc) which contain numerical data (points, lines, polygons
Technical Research Grade: These resources may be generated from data collected by regional or municipal governments, First Nations, community-based monitoring groups or other sources. Resources which receive this classification meet the standards for Technical Research Grade that have been developed by the Water Hub using community survey feedback, RISC guidelines and expert input. Although these resources have been determined to meet a certain standard of quality, they will still include a disclaimer as	

the Water Hub cannot guarantee the quality of any resource.	with accompanying water data)
Community Research Grade: These resources are generated from data collected by Citizen Science groups and community-based monitoring groups which may rely on volunteers and the public to provide the measurements collected in the dataset, and do not meet the criteria for Technical Research Grade. These datasets are still valuable, and may be very appropriate for use in some situations. An example of this data grade may include datasets collected by Selkirk College and College of the Rockies environmental programs, as well as other citizen science based projects such as Water Rangers. These resources will still undergo the QA/QC protocol for spatial and tabular, numerical data. These resources will include a disclaimer ensuring that any issues with the dataset are highlighted for the user.	
People and Perspectives: This could include written reports, images, descriptions of cultural values or other formats. Non-numerical resources will not be graded using the same standards as the spatial and numerical data discussed above, as they are assessed in a more qualitative manner. The intent is not to devalue or dismiss these resources, but to recognize that they cannot be consistently or fairly assessed using the same standards as numerical data. The following flowchart depicts the process used to determine the grade of a resource.	For non-numerical resources, including spatial resources which do not contain numerical data
Links: Data obtained through links to external databases or sources	Any type of data which is provided in the form of a link

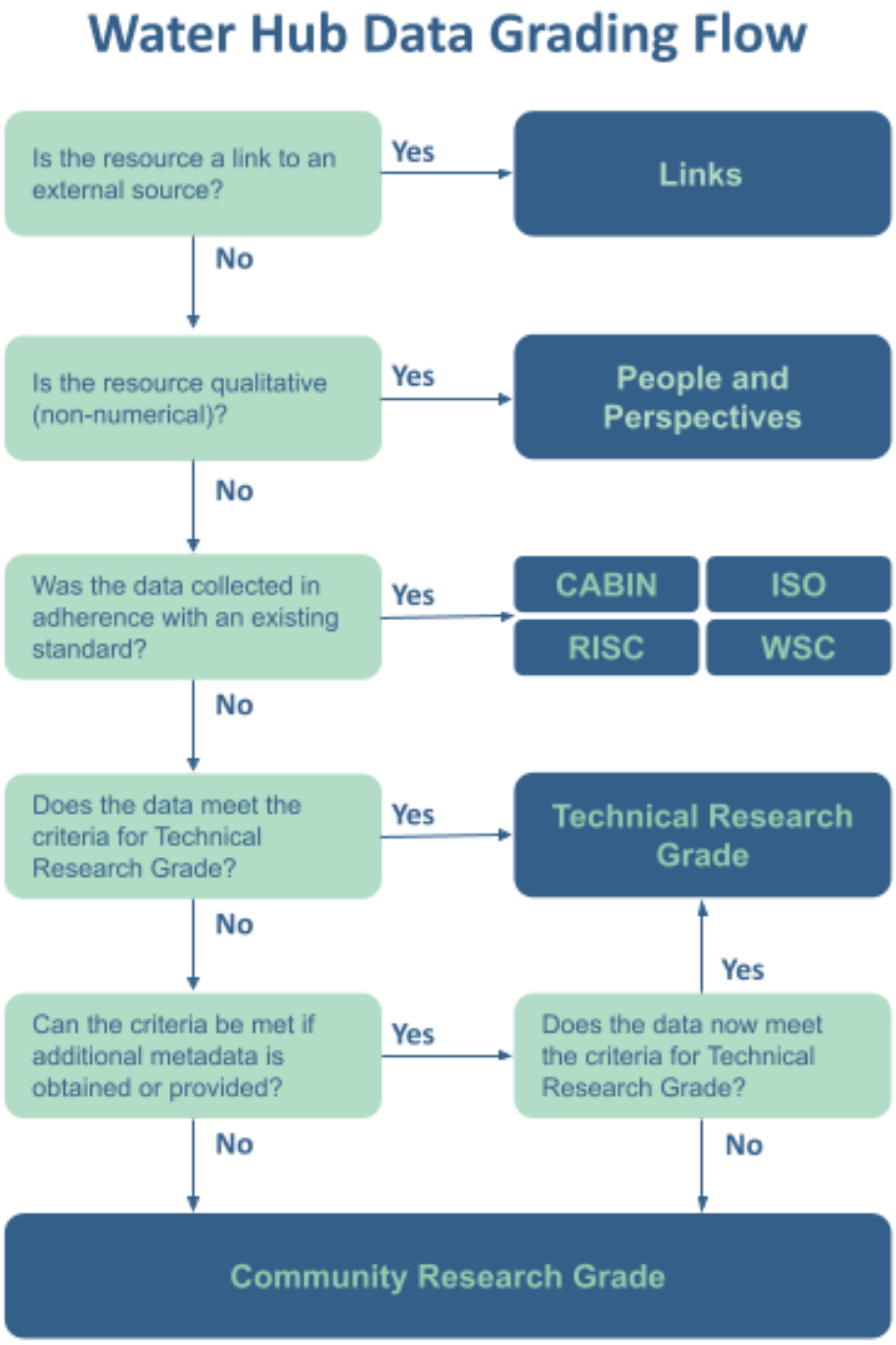

Figure 3: Flowchart illustrating the decision making process used to assign a Water Hub Data Grade to a resource.

Community Involvement and Feedback

The Water Hub focuses on being a place where everyone from every community within the Columbia Basin is able to access, use, and share water data for a variety of purposes. For this reason, the Water Hub will use a variety of approaches to encourage the involvement of all people whose lives are impacted by water. This means creating an outreach strategy which will bring the needs and values of First Nations, the public, government, academia, and industry together to create a cohesive approach to data monitoring in the Basin.

The QA/QC protocol will be adjusted as needed over time to ensure that it meets the needs of the community. By providing various and flexible approaches for the Quality Control and Quality Assurance of water data for all levels of monitoring, we hope to encompass the varied monitoring efforts, cultures, and ways to see our world that are present in the Basin.

The Water Hub Data Quality survey conducted in Winter 2021 was one example of the Water Hub engaging with the community. The feedback from this survey was used in the development of the Water Hub Data Grade classification system, and will also inform further outreach activities conducted by the Water Hub.

Responding to User Concerns

For all datasets, a reporting button will be added to the site where users can contact Water Hub staff if they have issues or concerns surrounding the quality of a dataset. If this occurs, Water Hub staff will assess the situation to try to determine the source of the issue. The data steward (as designated by the monitoring organization) may also be consulted. Depending on the source of the issue, the resource may be updated to correct an error (this will be noted in the Data Processing metadata field), an explanatory note or disclaimer may be added, or the resource may be removed from the site.

Possible Future Development of QA/QC

Promoting a Consistent Approach to Monitoring

As the next step to ensure that data meets the required standards of the community, in association with the Collaborative, the Water Hub will begin exploring the different methods currently used for monitoring data, and attempting to promote the use of standardized methods across the monitoring community.

This may include developing Water Hub Methodology which would be a more user-friendly version of Provincial standard methods, and reflects current literature and research. Organizations not currently using an existing standard would be encouraged to adopt this methodology.

For organizations currently collecting data according to an existing standard, the Water Hub may begin to request evidence demonstrating adherence to the standard (site visits, pictures, documentation etc.)

QA/QC for Data Hosted by the Water Hub, Uploaded by Users

In the future, monitoring organizations may begin to upload data without the direct oversight of Water Hub staff. This will only be permitted once an individual has received sufficient guidance or training on this process.

User uploads will not begin occurring until Water Hub staff are confident that the internal QA/QC and data management procedures are functioning effectively and that any issues with the process have been resolved. At that stage, it will be possible to develop a more thorough procedure and policy for user uploads, and this section will be updated at that time.

Users will select the Not Reviewed option for the Water Hub Data Grade field. The Water Hub Team, or other designated individuals, will conduct QA/QC on these Resources and assign a grade.

Database Management Roles

Applied Technology and Innovation Manager

Santiago Botero

Contact information:

santiago@livinglakescanada.ca

647-622-0166

Database and Community Engagement Coordinator

Paige Thurston

Contact Information:

paige@livinglakescanada.ca

5.0 Policies

The primary source for information about the policies of the Water Hub will be the Terms of Use that are posted on the Water Hub site.

Data Sharing Agreements

Each organization that is submitting or uploading data to the Water Hub will complete a Data Sharing Agreement. A template for this has been developed, which may be modified upon request of the Data Owner.

Private Data

The ability to create Private datasets may be granted to groups and datasets whose data contains any cultural, archeological, or otherwise sensitive information (identifying data). The metadata for these datasets will still be public on the Water Hub, but access to the actual data would only be granted in adherence with the specific organization's requirements.

Additional exceptions to free, unrestricted data access:

- Where local and Traditional Knowledge is used, Knowledge holders rights will not be compromised.
- If data release may cause harm to people or the environment, specific aspects can be kept protected or generalized to a broader geographic location.
- Where pre-existing data already has restrictions.

Hosted Data Policy

The CB Water Hub and LLC is dedicated to factual information and reserves the right to decline to share resources, or remove what is deemed controversial or incorrect content based on review by the Management Team, and Board of Directors.

Back Ups/ Archiving Procedures

The Water Hub aims to ensure that no data is ever lost due to technical issues. When relying on technology, it is always good practice to have a contingency plan. In order to ensure there is no data loss, every dataset will be backed up in a NAS (Network-Attached Storage) located in the Living Lakes office in Nelson, BC. This will also act as a backup for all the documents for the Water Hub and future Collaborative. data owners/uploaders can opt out of this, although it is highly encouraged to show that they have their own proper backup and archiving practices before opting out.

The stages of data management where backup is required are described in the Water Hub Data Management Flow.

Licenses

The Columbia Basin Water Hub allows for data uploaders to choose from a variety of different licenses that allow for proper accreditation of datasets. The list below provides the different licenses available when uploading data to the Water Hub.

- Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC-BY-NC-SA-4.0)

- Creative Commons Attribution ShareAlike International (CC-BY-SA-4.0)
- CC0 1.0 Universal (CC0 1.0) Public Domain Dedication (CC0 1.0)
- Open Government Licence - Canada (OGL-Canada-2.0)
- Open Government Licence - British Columba (OGL-BC)

6.0 Information Sharing

The process of transforming data into knowledge is a driving force behind the Water Hub's purpose. We aim to create analysis of water quality and quantity in the shape of reports and other formats within a defined interval.

This analysis will be guided with the scientific approach that was identified in the online workshop *Expanding Water Monitoring Within Canada's Upper Columbia Basi*n hosted by Living Lakes and led by Dr. Martin Carver and Greg Utzig. The proceedings can be found here.

Several avenues have been identified through which we can provide the public with information regarding their watershed's health.

We can use social media to highlight different community issues around the Basin, and how data has been a part of the solution.

Another approach is to create monthly reports for different areas of the Basin as stated in Paragraph 1, and provide them to communities free of charge. This initiative would encourage people to look into the issues affecting their watersheds.

Open source software and application development through platforms such as GitHub is another way in which we can turn our data into knowledge. These open collaboration tools can be used to develop water balance models, climate change tools and predictions, and other useful community focused applications. This will encourage innovation within communities, but it will also work as a bridge between the conservation and technology sectors.

7.0 Capacity Building

The Columbia Basin Water Hub is also going to be a place where people from across the Basin can receive training in data management, collection, analysis, GIS, and sharing. We hope to provide certificates confirming people have gone through our training program. This still has to be developed and could be a source of income for the Water Hub.

Training and technical support is currently being provided informally for community stewardship organizations. This includes support completing data management plans, uploading data, formatting data, and providing training on how to best upload and download data. The development of a small course would be ideal for providing a standardized way to certify people who can carry out their own QA/QC and upload their own datasets as “Water Hub Certified.”