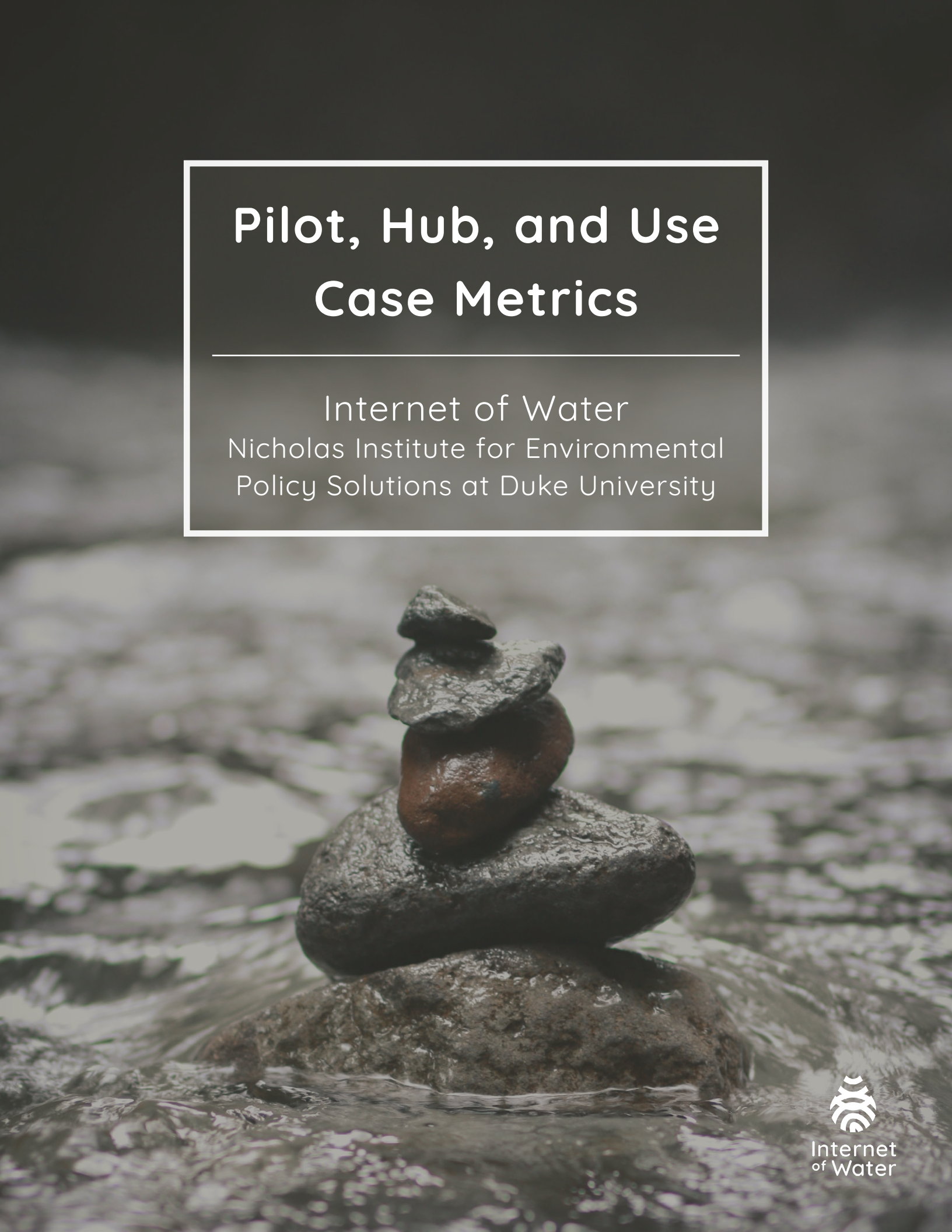
Pilot, Hub, and Use Case Metrics

Internet of Water
Nicholas Institute for Environmental Policy Solutions at Duke University



Internet of Water

Contents

INTRODUCTION

This guidance document is intended to help pilots, hubs, and data producers evaluate their efforts to promote water data. The guidance is designed to create frameworks for evaluation and develop assessment metrics. The guidance targets different stages of effort (e.g., stakeholder outreach, pilot design, and output evaluation, among others).

Evaluation is a systematic way to identify areas of improvement and communicate the effectiveness, impact, and value of an effort (whether a project or a program). While evaluation methodologies can be complex processes, there are a number of constructive approaches that provide useful information with less elaborate frameworks. Regardless of how simplistic or complex the evaluation framework, fully integrating evaluation metrics into an effort is necessary to understand the resulting impact and whether the goals of an effort have been accomplished.

Understanding and applying evaluation frameworks can encourage effective data management strategies, improve data infrastructures, and demonstrate the results of these investments. The purpose of this guidance is to:

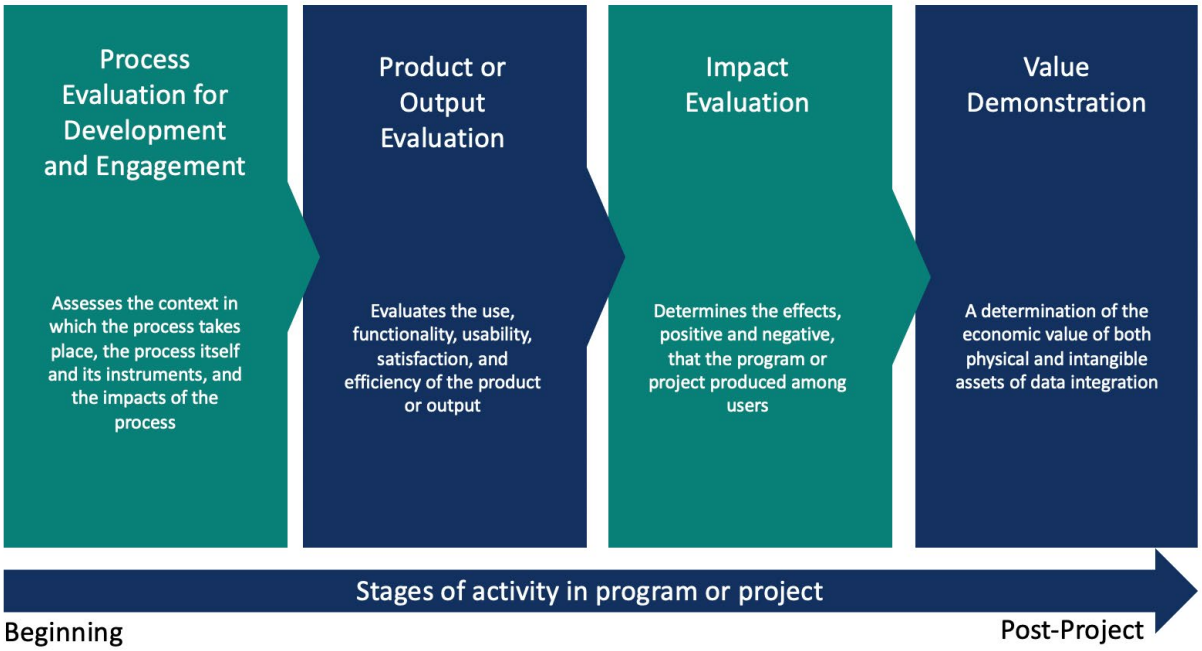
- Summarize the elements of evaluation
- Provide a framework for conducting effective evaluation(s)
- Clarify and provide examples of methods and metrics for evaluation

Given that evaluation and their measures are unique to each effort, rather than prescribe universal evaluations, this guide walks through the process of establishing a framework, and highlights a selection of the more common assessment metrics.

The metrics offered here are organized by the stage of the effort in which they are measured. There are four major classes of evaluation:

1. Process: assesses the context in which the stakeholder engagement place, the process of engagement and its instruments, and the impacts of the stakeholder process
2. Product: evaluates the use, functionality, usability, satisfaction, and efficiency of the product or output
3. Impact: determines the effects, positive and negative, that the effort produced among users
4. Value: determines the economic value of both physical and intangible assets of data integration

Integrating these classes into an evaluation framework and applying the framework to understand, improve, and communicate effort outcomes make up a comprehensive effort review. In turn, the information gained from a comprehensive review can be used to support decision-making, inform policy, and develop sound strategies for integrated water data management.



SECTION 1: PROCESS EVALUATION FOR STAKEHOLDER DEVELOPMENT AND ENGAGEMENT

Given that fragmented water data start with fragmented institutions, convening a variety of stakeholders is often a necessary first step to improving water data integration. Convenings around pilots and use cases where participants benefit from the work and realize the value of integrated water data at the same time are useful.

The purpose of process evaluation is to measure three aspects of stakeholder development and engagement:

A. The **context** of engagement (meetings, phone calls, surveys, online, etc.).
B. What activities **(approach)** make up the engagement..
C. Did the activities achieve the goals and what **impact** did they have.

WHY SHOULD YOU EVALUATE ENGAGEMENT?

Evaluating stakeholder engagement can provide valuable insight for future development and collaborative efforts. Stakeholders can provide input on engagement design, which can be used to refine the effort in ways that ensure sustainability. When stakeholders are

fully engaged in evaluation activities , they contribute substantively to the overall effort and become advocates in ways that ensure continued success. A thoughtful evaluation can reveal potential pitfalls in the effort, spell out how the effort can be improved, and clearly illustrate the potential of the effort. These lessons learned can be shared with higher-level decision makers to advocate for policy or change in current practices in ways that can be more effective than anecdotal examples.

Since engagement is about human experience, we recommend a survey approach to understand how individuals experienced engagement. Table 1 represents a sample survey that measures stakeholder engagement , linking the above questions to the objectives of the development and engagement evaluation outlined above (A. Context; B. Approach; C. Impacts):

Table 1: Sample survey for measuring the engagement process

Aspect Measured	Sample Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Context	I would like to have more input regarding the engagement sessions (locations, times, activities).					
Context and Approach	I feel I can talk openly and honestly about the issues facing my [community or agency] during the engagements.					
Context and Approach	My opinion is listened to and considered by the leaders of the engagement session.					
Approach	I am generally satisfied with the activities and progress of the partnership during the [time of effort].					
Approach	I have a sense of ownership in the [product or output].					
Approach	I trust the leaders of the engagement sessions to provide accurate information on [effort].					

Approach	I feel that the engagement sessions are a good use of my time in that we accomplish very much during these meetings.					
Approach	I would like to have more input into the development of the [effort].					
Approach	The communications surrounding the engagement sessions are timely and effective.					
Impact	Because of my participation in the engagements, I can have a positive effect on my [community or agency].					
Impact	I have used information from these engagement sessions to inform planning and activities for [effort] in my [community or agency].					

SECTION 2: PRODUCT OR OUTPUT EVALUATION

Now that stakeholders are working together, how do you evaluate what is created from these engagements? How do you ensure that the product or output of engagement is successful and meeting expectations?

Product or output evaluation focuses on the user experience. This may be the use of a web portal or decision-support tool, for example. These evaluations are intended to inform developers of such tools on a range of use and outcome metrics. For example, can the user correctly use and interpret the information from the tool? Does the tool function in the manner in which it was intended? Did the tool result in user satisfaction and efficiency improvements? Ideally these evaluations feed into an interative or ongoing product development cycle. As developers make changes based on the initial evaluation, further evaluations inform the development of the product until both user and developer expectations are met.

As with evaluations on engagement processes, product evaluation is about user experience; therefore, a survey approach is recommended. The exception to this approach

is the incorporation of technologies such as eye-tracking, when available. Eye-tracking technology is a sensor-based technology that measures the engagement of a user with an electronic or web-based tool. This type of evaluation offers tangible measures of user experience and allows evaluators to compare the experience across different groups of users (for example by age, gender, etc.) Unfortunately, eye-tracking can be an expensive and cost-prohibitive method for user evaluation. Outside of technologies such as eye-tracking, surveys are the most reliable method to evaluate user experience.

It is important to note that there can be bias in surveys following a collaborative process. If participants personally like the leaders of the evaluation, they will often feel the need to give only positive feedback, missing out on the opportunity to convey constructive feedback. To counter this bias, it is helpful to use observation techniques. In this case, users would be given a set of tasks to complete using the tool under evaluation. Observers would circulate during this interaction and note common points of confusion or error among users such as length of time to navigate from one point to another on the tool, where places of confusion may occur, and misinterpretation of tool output.

Table 2: Product or Outcome Evaluation Categories

Category of Evaluation	Definition of Category	Explanatory or Sample Question
Use	Used in planning, resource allocation, or policy decision	Have you used newly available datasets to make decisions about resource allocation or revise or develop policy?
Functionality	The tool or product works in the way intended	Were you able to use the tool/platform to find previously unavailable data?
Usability	The user can successfully access or gain from the tool or product	Were you able to access and compare datasets using the tool/platform?
Satisfaction	Likelihood of future use or recommendation of the product or output	Do you believe you will use the tool/platform in the future or recommend others to use it?
Efficiency	Saved time or money from use of the product or output	Do you believe you have saved time by using the tool/platform?

A distinction important to note is that between *functionality* and *usability*.

> *Functionality*, as stated above measures the tool itself, without regard to human interaction. Does the map react when the user clicks on a particular point? Does the menu selection navigate to the appropriate page?
>
> This is distinct from *usability*, which focuses on the user interaction with the tool. Is the user able to understand the meaning of menu titles to navigate to the appropriate pages? Is it clear where users navigate to run a tool outcome?

These two components are critical to the other categories of use, satisfaction, and efficiency. For example, if a user cannot clearly interpret how to navigate site menus, the user will likely be unsatisfied with the product. Because of this connection, it is important to place significant effort on understanding functionality and usability.

SECTION 3: IMPACT EVALUATION

Success in developing a product or launching a new data hub or app is exciting! At this point the focus shifts toward the impact of the effort. Did all of your hard work result in improved outcomes? Is the product or output having the impact as was hoped? If not, where are areas of improvement?

The metrics below outline a framework for assessing impacts. Previous research shows that collaborative work between experts and stakeholders (as opposed to expert-only designs) is more likely to result in enhanced decision-making and the integration of science and data into the decision-making process, enhancing the overall impact of the effort. However, assessing the degree of impact of these collaborations necessitates targeted evaluation methods designed to measure the success of such partnerships. Table 3 below represents a summary of the potential indicators for impact, definition of these indicators, and sample or explanatory questions that can be used as metrics for these indicators. These are participant-focused measures that evaluate the impact from a participant's point of view. However, these are not the only types of impact evaluation that can or should be conducted.

The growing number of web-based products that result from collaborative partnerships requires the integration of metrics designed specifically to understand the impact of these kinds of products. Google Analytics is one of the most powerful and commonly used tools for such web-analytics. It is also a free tool, which makes it available to anyone who wishes to engage in such evaluation. Table 4 outlines the most common and helpful Google Analytics metrics for evaluating the impact of a website or web-based tool. Like the product or output evaluation in Section 2, these types of metrics can identify areas of improvement to enhance the user experience thereby improving the impact of a website or web-based tool.

Table 3: Participant-focused measures of impact

Category of Evaluation	Definition of Category	Explanatory or Sample Question
Instrumental	User finds out what to do and how to do something or gains new skills	Have participants learned to use new technologies and methods to collect, manage, and analyze data?

Factual	The provision of precise data	Does the user have a better understanding of a particular dataset, where and how to find the dataset, and how to download or obtain the data from that dataset?
Conformational	Previous information was verified	Are the results of the data in line with what the participant or user expected? Do the data show what was expected or do the data contradict commonly held beliefs or notions about the dataset?
Projective	User gained better understanding of possible future scenarios	Is the participant or user able to use the dataset and/or tool for future planning or to understand and identify potential future issues?
Motivational	Encouraged someone to keep going on search for more information	Does the tool or product encourage the user or participant to explore other datasets with which he or she would normally not engage?
Personal or Political	Helped person gain control of a situation or avoid a bad situation	Did having access to a dataset allow you to challenge a policy or decision or avoid a bad decision?
Enlightenment	Self-perception to be better informed about an issue	Do participants and users feel better informed about data standards, best data practices, and basic data concepts?
Problem Understanding	More specific than enlightenment, better comprehension of particular problems	Do participants and users have a greater understanding of how integrated data can support sustainable water management?

Table 4: Product-focused analytics

Category of Evaluation	Definition of Category	Explanation
Unique Page Visits	The number of sessions during which a page is viewed one or more times	This category differs from simple page views in that it takes into account the same user going to a single page multiple times in a session. For example, navigating on and off a home page. Unique page views is an indicator of strong content and an engaging page or tool.
Bounce Rate	Number of users that enter website only to leave after viewing that single page	A high bounce rate indicates little engagement in the webpage or web-based tool. This in an indicator that work needs to be done to engage users or make their experience with a webpage or web-based tool more satisfactory.

Acquisition Overview	Displays traffic sources, such as whether a person navigates to the website via direct hit, referral, social media, or an organic search	Analytics overview is a key set of metrics that offers a snapshot of a webpage or web-based tool. This set of metrics offers a high-level set of metrics of daily interactions with a webpage or web-based tool. It is useful to monitor daily activity. T
Social Overview	Measures the impact of social media on the website or web-based tool	It is helpful to understand the link between social media and a webpage or web-based tool. Do you see an increase in use after a social media post about a web-based tool? This helps shape communications activity around your effort and indicates the impact of social media outreach on the use of a webpage or web-based tool.
Percent New Sessions	Percent of overall sessions made by new users	This is an indicator of growth, an important component of impact. Is your audience or user group for the webpage or web-based tool increasing over time?
Average Session Duration	The average amount of time of each session	The longer time a user spends on a page, the higher the engagement likely is between the user and the page. This is an indication of the quality of the engagement on your webpage or web-based tool.
Device Usage	The types of devices through which users access the webpage or web-based tool	Webpages and web-based tools should be responsive to different user environments. Whether a user interacts with the webpage or web-based tool via a desktop computer, mobile device, or tablet, the goal is a positive user experience. This metric will show the types of devices through which users access the webpage or tool, allowing for the opportunity to optimize the product for the user experience. If you have most users accessing your webpage or tool from a tablet, it is beneficial to ensure the webpage or tool is responsive to a tablet environment.

Other types of indicators for impact can include:

- Feedback on effectiveness or benefit of effort, tool, portal, etc..
- Observable changes in attitudes, behavior, skills, or knowledge among participants and users
- Person hours saved by use of the product output

And there are likely many more. Evaluating impact requires an integrated approach to evaluation, likely incorporating varying methods to answer a variety of questions. It is important to note the distinction between evaluating impacts and monitoring outcomes.

Monitoring outcomes refers to a description of factual evidence or observations, and using this description to attribute observed outcomes to a specific intervention.

Impact evaluation determines the success of goals of a effort and establishes whether this effort had effect on individuals, communities, agencies, decision-makers, stakeholders, etc. Impact evaluation is focused on lessons learned as well as identifying effort efficacy.

This type of evaluation can be used in decision-making to determine future support, upscaling, or expansion of a effort. Well-designed impact evaluation can answer questions such as: which parts of the effort work and which do not? What policy-relevant information should be incorporated into a redesign of the effort? How should future efforts be designed? Why and how does the current effort work, or if is not working, why not?

SECTION 4: VALUE DEMONSTRATION

Being able to show the value of a water data effort is important in getting resources in the future. Value can be internal time savings, internal or external water savings, and opportunity costs associated with integrated water data. Incorporating value demonstration into your overall evaluation framework helps communicate to others and advocate for support.

The value of water data has not been well documented, quantified, or communicated. The absence of such a framework for valuing water data requires those who wish to determine such value rely on other methods of valuation found in other fields.

Table 5 represents a summary of the more useful methods to value water data for hubs, producers, and users. Listed for each method is a definition, the intended audience, and suggestions for the types of data needed to carry out the method of valuation. These methods are further explained and modeled here: https://internetofwater.org/data-stories/valuing-data/#Approaches

Table 5: Valuation Methods

Name of Method	Definition of Method	Intended Audience
Market	Assesses user's willingness to pay for a product or service. The data for this method is commonly estimate through consumer surveys and current market prices for goods and services.	Producers of Data

Modified Historical Cost	Treats data as an asset whose value is at least equivalent to the cost of data collection. Data costs include labor, equipment, infrastructure, up-front capital costs to install equipment, ongoing operation and maintenance costs, costs to store, process, and use the data.	Producers of Data
Business Model Maturity Index	A top-down approach that assesses the value of data based on their relative contribution to a final outcome, using use cases to evaluate better informed decision-making. Identified outcomes are compared with potential impact using metrics such as time savings, cost savings, water savings, lives saved, etc. This is compared to implementation costs of the use case and the potential impact accounted for over the life-time of the effort.	Users of Data
Decision-Based Valuation	Estimates the relative contribution of data while accounting for data attributes (quality and frequency of collection, for example), relative to the decision being made. Identified outcomes are compared with potential impact using metrics such as time savings, cost savings, water savings, lives saved, etc. Use cases are developed, and estimated costs and impacts are calculated. These are adjusted based on how fit-for-purpose data are to inform decisions. For example, frequency compares how often the data must be collected to inform decision-making; accuracy is compared with the required accuracy for the use case; an quality modifiers are calculated based on the sum of the frequency and accuracy scores. Additionally, costs are calculated based on labor, hardware, software, infrastructure costs and any contractual work conducted outside the organization.	Users of Data

Consumption-Based	An adjusted Modified Historical Cost Method, adjusted to estimate the value of data hubs, assuming data is received from different producers and shared to multiple users. Total annual expenditures or expenditures by dataset or individual data producers, adjusted to reflect the value of data relative to their use within the hub by scoring their usage (zero equals no usage and one equals maximum usage, for example). Additionally, value is adjusted for data quality.	Data Hubs
Keep Research Data Safe	Calculates the long-term costs of data hubs as well as the value of hubs to data users. Identify costs of hub such as yearly expenditures, costs by individual employees or datasets, and depreciate these costs and adjust for inflation. Benefits are calculated with direct and indirect impacts such as increased usage, decreased data discovery time, discarded or lost data inefficiency. In addition, benefits can be grouped into short-term (realized in less than 5 years) and long-term benefits (longer than 5 years to realize benefit), as well as if these benefits are internal to the hub or external for stakeholders.	Data Hubs

SECTION 5: TYPES OF EVALUATION TOOLS AND INDICATORS

There are many types of evaluation tools that can be employed when developing an evaluation framework. It can be difficult to determine which tool is best deployed for a specific type of evaluation. The type(s) of evaluation tool(s) should be informed by both the type of data output desired (quantitative or qualitative), the target audience for the evaluation, and the available capacity and resources. For example, it may not be practical to conduct individual participant interviews. In such cases, a focus group may provide needed feedback with much less time and organizational commitment.

Table 6: Types of evaluation tools

Evaluation Tool	Description of Tool	Type of Tool
Survey	A set of predetermined questions about topics that are answered by a target audience	Quantitative
Interview	A set of questions (may or may not be predetermined) about topics that are posed to a target audience and followed-up with additional questions	Qualitative
Knowledge or Skills Tests	A set of questions that determine the level of knowledge or skills among participants	Quantitative
Focus Group	Group discussions with a relatively small number of selected people about certain questions	Qualitative
Evaluation Form	A set of questions that determine the participants' opinions, attitudes, and understanding once a effort activity is complete	Quantitative
Journal Recordings	Self-reported, daily activities of users or participants	Qualitative
Activity Log	Self-reporting of daily activities	Quantitative
Anecdotal Records	Stories and narratives about a effort, participant experience, or event described by participants	Qualitative

CONCLUSION

Regardless of the methods or metrics used to assess or evaluate a effort, it is essential that evaluation be integrated into the overall planning and design of the effort. Far too often, evaluation is an add-on at the end of a effort, limiting the type and robustness of the evaluation that can be conducted. By considering and planning for evaluation from the outset of the effort, success and progress can be fully demonstrated because the data required for such evaluation is collected and compiled in a manner most helpful. The results of evaluation can enhance communication about the effort; be used to improve the process, product, and impact of a effort; and advocate for support when needed.

This guide is not intended to be a comprehensive guide to all evaluation. Instead, this guide is specifically designed to address the context of pilot, hub, and use cases around integrated water data initiatives. It is recommended that users of this guide also seek guidance from the variety of resources available regarding evaluation to determine the metrics and methods that best meet organization goals and objectives. To the extent possible, this guide provides a framework for establishing an evaluation approach, examples of how to carry out evaluation activities, and sample metrics to measure effort goals and outcomes.