ITRC has developed a series of fact sheets that summarizes the latest science, engineering, and technologies regarding environmental data management (EDM) best practices. This white paper describes:

- top ten things to consider when planning public communication
- principles and concepts of public communication and data accessibility, including the findable, accessible, interoperable, and reusable (FAIR) data principles, tools for communication, accessibility both on the web and in the physical world, and the use of plain English
- a discussion of stakeholder identification and characterization
- guidance on tools to use for stakeholder communication, and the importance of building and executing a stakeholder communication plan
- a look at how public communication shapes database design

Additional information related to public communication is provided in the ITRC fact sheets under the Data Management Planning, Geospatial Data, and Traditional Ecological Knowledge topic areas.

1 INTRODUCTION

Detailed and standardized environmental data management (EDM) is essential for regulatory agencies to make well-informed decisions to improve and protect human health and the environment. Communicating why and how these decisions were made to all stakeholders (internal, external, and the general public) and providing appropriate access to the data and metadata are equally important for successful implementation and serving public needs. Some examples of public communication are:

- members of the public noticing the closure of beaches due to sewage overflow causing high bacteria levels in coastal waters
- real-time monitoring of fire-related particulates in the air
- documented contamination in soils and groundwater that may exist near residential properties or drinking water wells
- up-to-date availability of campsites at your favorite regional, state, or national park
- posting fishing advisories

With transparent, understandable, and widespread outreach, public health is enhanced and damage to people, property, and the environment may be decreased or avoided.

The data lifecycle (see Data Lifecycle Fact Sheet) is a conceptual path describing how data are collected, stored, and retained. Its components include planning for collection and end-use objectives, acquiring data, processing, and maintaining data for format and usability, publishing and sharing data, and retaining data. It is important to determine the communication objectives and any end-user call-to-action for your environmental data during the *planning* phase of the data lifecycle. These objectives will directly impact how the data are *acquired* and *published/shared* during the duration of the project and beyond. Figure 1 provides a visual representation of the data lifecycle with shadows indicating the three steps (plan, acquire, and publish/share) in the process where public communication considerations are most applicable.

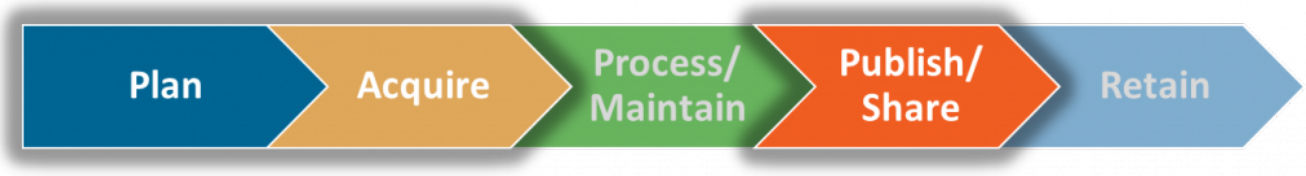


Figure 1. Phases in the data lifecycle where public communication considerations are most applicable.

To effectively communicate, regulators need to define the objectives of the intended communication, identify stakeholders, define key messages, and identify communication methods to best share the information (Jarreau et al. 2017). In the context of this document the audience is the population of the United States: a person, group, or organization that is affected, potentially affected, or has any interest in a project or a project's outcome, either directly or indirectly. A stakeholder is a person or organization, including responsible parties, or an agency that is invested in or impacted by a situation or issue. A user is a person, organization, agency, or other governmental entity that directly interacts with products or projects. Some

users are also stakeholders, and all are part of the audience. Additionally, users may be internal (part of the team within the organization that owns/maintains the data) or external (others outside of the organization that owns/maintains the data).

Stakeholders can be categorized into two camps—internal and external. Internal stakeholders are team members, different departments, or management within an agency or organization; please see the Data Access, Sharing, and Security Fact Sheet for a more detailed discussion of internal stakeholders and how they fit into the data lifecycle. This document focuses on communicating environmental data with **external stakeholders**—those that exist outside of the agency or organization.

This white paper covers principles and concepts of public communication, identifying groups of stakeholders for targeted communication and data sharing, the importance of developing and implementing public communication plans, and the design and use of public communication data platforms.

Top Ten Things to Consider When Planning Public Communication

1. *What story do you want to tell?*
    - *Is there a risk that needs to be communicated?*
    - *Will people need to take action based on the level of risk?*
    - *Who is it impacting?*
    - *Where and how will data be collected?*
    - *Is there a behavior that you are aiming to change?*
2. *Are there political considerations from your management/agency that need to be considered?*
    - *These might not match with what the governor or other entity wants to share and may impact when/what you can share about the status of a situation.*
3. *Who will want to use the data, including those outside of the intended audience, and how will they use it?*
4. *What data presentation formats are most appropriate for the different user groups?*
5. *What are the information needs of the local community regarding the topic?*
6. *What is the community's history with…:*
    - *…the agency?*
    - *…the specific environmental issue?*
    - *…local politics?*
    - *…local values?*
7. *What is the best way to frame information to make it relatable to your audience?*
    - *Use metaphors and analogies to make complex topics more understandable*
    - *Economics*
    - *Health benefits*
    - *Avoiding waste*
    - *Value-based actions*
8. *What languages are spoken in the community?*
    - *Find people within the agency who are fluent in the community language(s) to draft outreach documents.*
9. *What is the technology level of your audience?*
    - *Internet access and how to use the internet*
    - *Understanding of common software*
10. *Proofreading! The more eyes on a document before it is published, the better.*
    - *Use a relatable tone in the deliverable*
    - *Check for spelling and grammar errors*
    - *Limit use of acronyms and abbreviations*
    - *Ensure message clarity*

2 PUBLIC COMMUNICATION PRINCIPLES AND CONCEPTS

2.1 Principles and Concepts of Communication

There are many principles and concepts that may be used to effectively communicate and share data with the stakeholders, users, and the general public. Important principles include proactive data transparency whenever possible, adhering to data sharing agreements that may not allow specific data to be publicly shared, making the message personal to the users, and providing local context.

Data is a valuable public asset. Maintaining its value includes making the data widely available in usable formats with intact metadata and proactively making data publicly available and findable. This helps to relieve some of the work associated with fulfilling public records or data requests, alleviating staff stress, time away from routine tasks, and the cost of fulfilling individual requests. Data transparency, including the source, quality, and impact to the public, helps build trust between government agencies and impacted communities. Additionally, it is important for data to be available beyond the scope or life of the original project. This is motivating for the data collectors to know that their valuable data may be used repeatedly in future projects.

However, as mentioned above, not all data can be made publicly available. For example, traditional ecological knowledge data collected by or through consultation with Indigenous communities and governments, homesteaders, etc., is considered the private property of those entities and could contain knowledge sacred to those groups. Data sharing agreements should be considered when working with these groups and governments to ensure their data are treated with respect and privacy. The CARE principles developed by the Global Indigenous Data Alliance encourage data managers to consider both people and the purpose in their work. In summary, the CARE principles are:

- ***c****ollective benefit*: Design data ecosystems to function in ways that enable Indigenous Peoples to derive benefit from the data.
- ***a****uthority to control*: Recognize Indigenous People's rights and interests in Indigenous data and empower their authority to control it.
- ***r****esponsibility*: Share how Indigenous data are used to support Indigenous Peoples' self-determination and collective benefit.
- ***e****thics*: The primary concern at all stages of the data lifecycle and across the data ecosystem should be Indigenous People's rights and well-being.

For more information on traditional ecological knowledge, please see the Traditional Ecological Knowledge fact sheets. For further discussion on data privacy and security, please see the Data Access, Sharing, and Security Fact Sheet.

Depending on the type of information you are sharing or the desired outcome, focusing on local context and using human stories to make the message more personal can yield better understanding of the environmental data and conclusions. Targeting audience concerns and values, using environmental psychology–based communication principles, establishing a connection between your audience and the local environment, and personalizing messages using real-life experiences are all powerful tools (Jarreau et al. 2017). Think of asking questions such as: How do residents feel? Do they feel ignored? What is their current understanding of the environmental issue and what information would help increase their awareness?

2.2 FAIR Data Principles

2.2.1 Introduction to FAIR Data Principles

In 2016, a diverse set of stakeholders published the "FAIR Guiding Principles for scientific data management and stewardship" in *Scientific Data* (Wilkinson et al. 2016). The FAIR data principles put emphasis on enhancing data reuse by increasing the ability of machines to find and use data as people increasingly rely on computational support to interact with data. The authors of the article provided guidelines to improve the **F**indability, **A**ccessibility, **I**nteroperability, and **R**euse of digital data assets. Since publication, these principles have been broadly explored, embraced, and promoted across many scientific data communities.

The FAIR data principles apply to data, metadata, and supporting infrastructure and encourage the findability and reuse of data in ways that reach as many people as possible. More information on the FAIR data principles is available at the Enabling FAIR Data Project and GO FAIR Initiative. In the sections below, the use of "(meta)data" implies that the statement may apply to data or metadata.

2.2.1.1 Findable

Findable data principles emphasize that data and supporting metadata must be findable to both people and machines to be reusable; making data more findable allows users of data to easily access data through search engines or other means. Findable principles are:

- (Meta)data are assigned a globally unique and persistent identifier.
- Data are described with rich metadata.
- Metadata clearly and explicitly include the identifier of the data they describe.
- (Meta)data are registered or indexed in a searchable resource.

2.2.1.2 Accessible

Accessible data principles emphasize that data should be available with as few barriers as possible; once data users find their desired data, they know how the data can be accessed. Accessible principles are:

- (Meta)data are retrievable by their identifier using a standardized communications protocol.
- The protocol is open, free, and universally implementable.
- The protocol allows for an authentication and authorization procedure, where necessary.
- Metadata are accessible, even when the data are no longer available.

2.2.1.3 Interoperable

Interoperable data principles emphasize that project data need to be easily integrated with other data or applications for analysis, storage, and processing. Interoperable principles are:

- (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- (Meta)data use standards that follow FAIR principles.
- (Meta)data include qualified references to other (meta)data.

2.2.1.4 Reusable

The above principles allow data to be reused. Reusable data principles emphasize that data and metadata should be described so that they can be used beyond the life of the original data collection effort. Reusable principles are:

- (Meta)data are richly described with a plurality of accurate and relevant attributes.
- (Meta)data are released with a clear and accessible data usage license.
- (Meta)data are associated with detailed provenance.
- (Meta)data meet domain-relevant community standards.

2.2.2 Incorporation of FAIR Data Principles into a Data Management Plan

FAIR data principles should be considered during the planning phase of a project and incorporated into the environmental data management system (EDMS) design as well as the data management plan. The following should be considered:

- Import metadata that can be clearly defined, documented, and archived for delivery.
- Define data standards early in the life of the project, allowing for (meta)data to be clearly defined and qualified, and incorporated into practices from field to publication.

Taking the time to clearly plan and document the execution strategy for FAIR data principles will ultimately save time in data delivery, improve the quality and context of the data, and make it easier to reuse the data in the future.

2.2.3 Budget and Timeline Considerations for FAIR Data Principles

Elevating data quality to meet the FAIR data principles may add cost and time to a project. However, the additional effort will allow more people and machines to find the valuable data that have been collected. Embracing FAIR data principles will

extend the use of data well beyond the life of the original project. Remember, other data users are an important stakeholder to consider during the planning phase of a project.

2.3 Plain Writing Act (Federal)

The Plain Writing Act of 2010 requires that federal agencies use clear language that the public can understand and use. Although the law specifically provides guidelines for federal agencies, these guidelines help to facilitate clear communication with the public for any public agency. Plain language is useful at all communication levels, from web content and social media, to technical reports. An extensive amount of training and resources is available to empower agencies to communicate in clear, plain language, and several groups and professional organizations serve as a community of practice for plain English communication.

The Plain Writing Act was developed to encourage government agencies to produce documents that are more effective and accountable to the public. The Plain Writing Act defines plain writing as "writing that is clear, concise, well-organized, and follows other best practices appropriate to the subject or field and intended audience" (Plain Writing Act of 2010, Section3(3)). The use of plain writing can be applied to documents such as regulations, policies, forms, social media, and reports. Using plain writing can be difficult when trying to present complex or jargon-filled scientific information. Suggestions to keep writing plain include:

- stating who the document applies to and what the stakeholder should do with the information presented
- avoiding jargon and using more words than needed
- obtaining ongoing feedback from stakeholders and the public
- avoiding acronyms

The benefits of plain writing include:

- improving the public's understanding of requirements
- assisting stakeholders with required compliance
- reducing errors
- increased public use of products and publications

Understanding your audience and the desired outcome of a document is key to plain writing. Due to a lack of resources or short timelines, authors will often try to address separate audiences with one overall document. Separate audiences have diverse levels of knowledge and understanding of a topic and mixing information in the same document can be confusing. Where possible, authors should address separate audiences independently.

Plain language emphasizes content and engages audiences with active voice and short sections with concrete, familiar words. Organize documents by using headings, lists, tables, and figures when possible. It is important to define uncommon or technical terms for the average reader. Avoid the use of acronyms and abbreviations when possible; instead, write out terms or use shortened phrases. In relation to data deliverables, use plain English in narrative descriptions that may accompany data, such as a narrative report from a laboratory that supplements analytical results.

2.4 Data Accessibility

When delivering information or data to the public, make sure it can be accessed by the target audience and stakeholders. Accessibility is a fairly broad term. For purposes of this section, we are using the term "accessibility" to reference how data and conclusions are brought to the public, and informing them about actions they can take to maintain their health and safety. Meaningfully consider the data consumer's frame of reference and circumstances. Convey information in a relatable manner; for example, economically, environmentally, or values-based (Corner, Shaw, and Clark 2018). Ultimately, data accessibility strives to ensure that the audience can understand, navigate, interact with, and take action to maintain their safety. For accessibility of data transferring from one user or location to another, please see the Data Exchange fact sheets.

Data accessibility factors to consider include:

- **Data format—**What format works best for the audience? Do they want raw relational data tables, or do we need to summarize the data and present it visually?

- **Platform access—**Can the audience access the proposed platform? Do they have internet access or modern equipment to understand and interpret the data in the way it is presented?
- **Equipment/software requirements—**Does the audience have the equipment or software to access the data as presented?
- **Technological capability—**Does the audience have the technical skill set to access the data as presented?
- **Language—**What language does the audience speak and do the data need to be translated? What is their level of English proficiency?
- **Culture—**Are we presenting data in a way that is relatable to the consumer's specific cultural background? Could the audience perceive the data as offensive?
- **Trust**—What is the nature of the relationship between the data consumers and the regulatory agency? If there isn't trust, what steps can you take to build trust?

In addition to the factors discussed above, the federal government and some states have laws or regulations related to accessibility. In this context the term "accessibility" refers to ensuring that information technology is accessible to people with disabilities. The relevant federal rules in Section 508 of the Rehabilitation Act of 1973 can be found on the IT Accessibility Laws and Policies website. This type of accessibility generally refers to public-facing websites but can also be required for documents and other forms of governmental data. Considering digital accessibility for those with disabilities generally involves understanding how users of different abilities will access content and often includes preparing for content to be consumed by tools such as screen readers. Companies such as Microsoft and Adobe have developed tools to check for content accessibility. It can be particularly challenging to adequately describe visualizations (figures, maps, charts, etc.) via textual descriptions and to provide access to raw data tables, as asked for by current web accessibility guidelines. The paper Rich Screen Reader Experiences for Accessible Data Visualizations (Zong et al. 2022) presents examples of design structure, navigation, and description that create screen reader experiences that help users conceptualize data spatially, focus on different levels of data granularity, and control their data analysis process.

Additional reading—Additional information about web contact accessibility can be found on the World Wide Web Consortium's Web Accessibility Initiative (WAI) website. The federal General Services Administration's (GSA) accessibility website also provides information about creating accessible documents (Create Accessible Documents | Section508.gov).

Advice for state regulators—Be aware of your own state laws or regulations related to accessibility requirements when developing your communication plan and developing your end products.

2.5 Risk Communication and Conveying Uncertainty

Effectively managing risk, along with clear communication of risks to the public, improves the public's understanding of ongoing uncertainty and how to proceed during high-stress situations, such as responding to or evacuating from natural disasters (for example, during hurricanes, wildfires, tornadoes), or responding to an unauthorized release of hazardous materials resulting from a truck or train accident. Several useful references and well-developed toolkits for responding to emergencies and communicating risk to the public are presented in the list below:

- ITRC's Risk Communication Toolkit
- SALT Approach developed by the U.S. Environmental Protection Agency (U.S. EPA) and focusing on strategy, action, learning, and tools
- Crisis and Emergency Risk Communication Toolkit developed by the California Department of Public Health

As with other types of communication involving environmental data, successful risk communication requires identifying and characterizing your audience, and how best to inform them about environmental risk and what actions to take to keep themselves safe.

Uncertainty about data that are analyzed to make recommendations or create action levels must be shared as data and conclusions are reported out to the public. The level of uncertainty should be described in an understandable way using plain English. It may be presented like a weather forecast: rain is forecast, but maybe only a 50 percent chance. After releasing a communication, be prepared to answer follow-up questions about how accurate data or analysis of the data is. During the planning phase when considering how to communication to the public, also consider how to document and explain potential uncertainties. This proactive transparency will help build trust with the audience and provide for a better understanding of the information that is communicated. For further discussion of conveying uncertainty in data, please see

the Data Quality Overview Fact Sheet, Table 1 of the Using Data Quality Dimensions to Assess and Manage Data Quality Fact Sheet, and EDMS White Paper.

3 STAKEHOLDER COMMUNICATION TOOLS

3.1 Public Communication Methods and Platforms

A public environmental data communication platform is a great way to share data and actionable information based on the data with the users, stakeholders, and the general public. In broad terms, the platform consists of a data source(s) and one or more communication pathways funneling the data or information to the public. Typically, the data source is a database, and the communication pathway is a website (including interactive maps, tables, or query-driven downloads), printed material, email list serve, or a web services application programming interface (API), etc. A goal of all pathways is for them to go both ways: communicate out and be able to receive feedback and questions—always include contact information.

Below are some examples of communication pathways:

- email list serves, newspaper notifications, local workshops, door fliers, website, cellphone apps, signs (physical or digital—for example, Do Not Swim), etc.
- an agency's public communication director/officer's official channels
- fact sheets containing current project status and contact info
- newspaper notifications
- webinars
- social media
- posting through news outlets and media
- blog posts
- push/pull or (more formally) proactive/reactive communication
- a public environmental data communication platform (discussed below)

A public environmental data communication platform enables data to be analyzed, visualized, or interpreted. Agencies can make available raw data and communicate important findings that convey the importance of environmental data to stakeholders. An effective public data communication platform communicates information with specific outcomes in mind. Typical outcomes might be to change public behavior to keep people safe, inform the public of environmental risk, or provide a portal for data transfer/download to support an investigation.

Example outcomes of communicating environmental data to the public include:

- affect a beneficial change in short-term public behavior, such as
    - preventing swimmers from entering unsafe oceans or lakes
    - modifying people's daily activities based on air quality
    - preventing consumption of unsafe drinking water
- inform the public about environmental impacts and risks, including
    - groundwater/soil/sediment contamination
    - environmental risks limiting brownfield redevelopment options
    - rivers/lakes/aquifers not meeting beneficial use requirement
    - environmental factors affecting construction development
- communicate current and past environmental conditions for activities such as
    - understanding baseline conditions
    - researching long-term changes
    - phase 1 environmental studies
- monitor changes in environmental condition (improving/degrading) that might require action, including
    - permit compliance

- beneficial use
- identifying and studying emerging compounds of concern

Example platforms:

- AirNow: A U.S.-focused air quality web page presenting easy-to-understand graphics, maps, and key performance indicators
- CalEnviroScreen: Filterable online maps, data, and reports developed to help identify communities in California that are disproportionately burdened by multiple sources of pollution
- CA Safe to Swim: Helps Californians be aware of swimming advisories
- GeoTracker: California Water Boards' visualization and data management system for various site types that impact, or have the potential to impact, water quality in California, with emphasis on groundwater
- IQAir: Visually powerful maps and charts that allow the user to explore the air quality anywhere in the world
- Purple Air: A website map showing an internet of things network of sensors, many of which are maintained by citizen scientists
- Washington EIM: The Washington State Environmental Information Management System (EIM) containing environmental monitoring data collected by state scientists and partners. Use the EIM map or form to search, view, and download environmental monitoring data for air, water, soil, sediment, aquatic animals, and plants
- Minnesota Groundwater Contamination Atlas: A tool for learning about polluted sites around the state of Minnesota. Use the database and atlas to learn the story about contamination at a specific site and to download groundwater sampling data

An effective public data communication platform identifies and targets its intended stakeholders in development of map services, ease of access, and types of downloads that are made available. Anticipating the use cases for stakeholders can allow for designing the data management system for more efficient and effective communication. For example, many stakeholders may be impacted by an environmental issue but are not aware of or have access to an existing environmental data communication. In this case, knowledge of the stakeholder could lead to a more effective communication strategy. Please see the Data Governance, Geospatial Data, and Data Exchange fact sheets for more information on planning, visualization, and movement of data.

Example stakeholders include:

- general public
- consultants
- laboratories
- other state and local agencies
- other governmental bodies (federal or state government, tribal nations)
- citizen scientists

Public communication is not a one-size-fits-all process. Identifying groups of stakeholders and determining a variety of communication methods to reach unique groups goes a long way toward successful communication of environmental data. Evaluating available census data might allow one to better understand communities and better engage with the public and stakeholders. Making data readily available in formats that are usable for different use types, transparency, and understandability are major principles of effective communication.

3.2 Tools Supporting Communication

There are many types of digital tools and processes supporting data communication. These can be open data portals, such as the federal government's Data.gov, geospatial and map services, web services and APIs, and broad-reaching automated messaging services, such as Granicus.

Geospatial and map services can be static or interactive and come in the form of tables, figures, and maps. Complex geospatial databases may have web-based analytical tools, widgets to query specific data, and various layers to allow for multivariable parameter views. Design the interface with the general public in mind—make it easy to use and adaptable to a

user's physical limitations. Create how-to-use guides such as a virtual tour of the service.

Web services and APIs allow computers and applications to communicate with each other. These tools are a powerful way to allow end users to be programmatically retrieve, or "pull," data from a data set. Once configured, stakeholders can add web services and APIs to routine processes so data flow more smoothly. For example, a local weather reporter may use an API to pull data from a water management agency into her models, allowing seamless reporting of river conditions alongside precipitation information.

Data communication typically falls into either the "push" or "pull" data flow models (Figure 2). The communication of environmental data occurs between agencies and stakeholders or other users. A stakeholder "**pulls**" data when they request information from a database through a web query. Conversely, the **"push"** model happens where the data flow is initiated from the data system to the stakeholder. A representative "push" communication would be an email to a new property owner informing the property owner of environmental conditions. In a push mode, a "trigger" for the communication identifies the stakeholder and the context. In the property owner example, the trigger could be a county tax roll record showing a new property sale and the identification of the new property owner. New triggers are scanned for by the data system, which initiates the push communication.

Both push and pull communication have comparative benefits. In a pull model, the stakeholder is typically aware of the data system and brings user context to the session. The stakeholder would discover the data system based on public outreach or marketing. In a push model, the stakeholder would generally lack awareness of the data system. Their engagement with the communication would be increased by the context to the communication. For example, the new property owner could receive communication that identifies their property and property-specific environmental information.

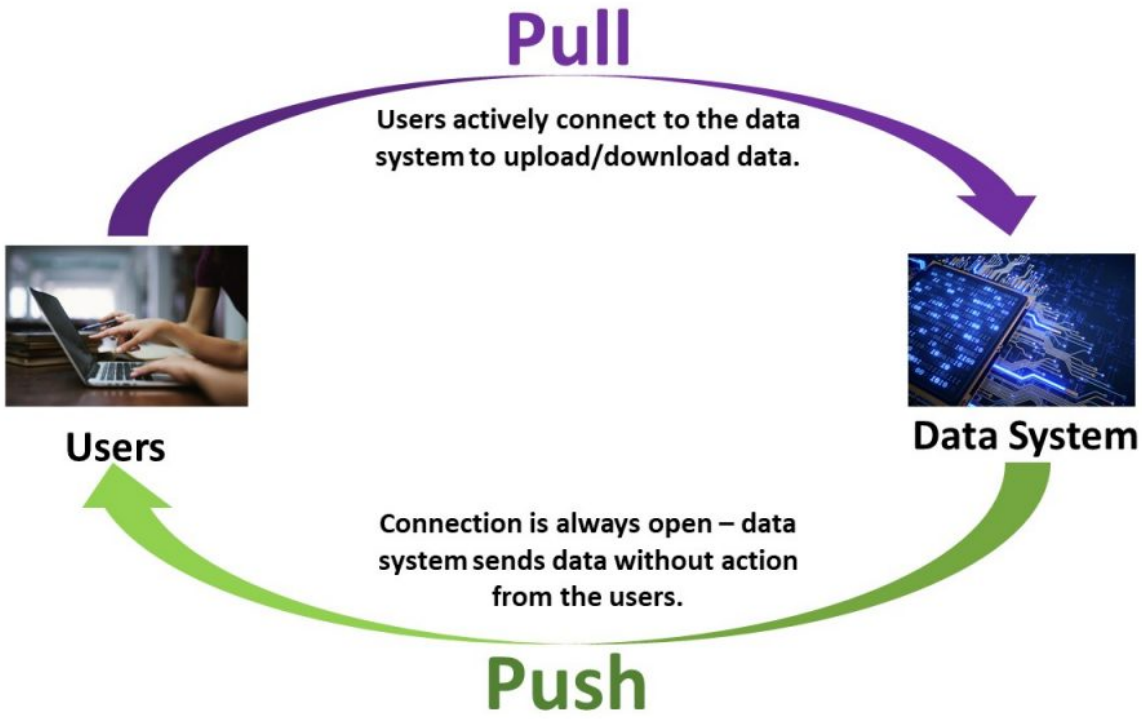


Figure 2. Data push/pull model.
For further information on tools organizing environmental data and supporting environmental data communication, please see the Geospatial Data and Data Exchange fact sheets and the EDMS White Paper.

3.3 Stakeholder Engagement Plan

Communication with stakeholders is an important part of a project, and a formal plan for engaging with them ensures that vital stakeholders' needs are met. Stakeholder engagement plans can vary in detail and complexity, but address the fundamental topics of:

- identifying stakeholders, whether active (for example, funder) or passive (for example, unknowing member of public)
- assessing the needs of each stakeholder and the tone needed for communication activities

- methods for successful communication

Metrics for successful engagement should take place early in the life of a project to be able to address stakeholder needs throughout the life of the project. Some stakeholders may want periodic updates on project status, or want preliminary data, in which case, communication frequency must be planned.

If communication is focused during an intensive time, such as at the end of a project, stakeholder engagement can be organized in a campaign-like event. With this style of outreach, the same stakeholders are targeted with various communication methods. Stakeholders are offered the same message through various lines of communication to enhance understanding, prevent fatigue from one medium, and reach diverse parts of a particular audience. Note that this sort of major outreach requires a lot of planning—be sure to start early to meet communication goals.

3.4 Stakeholder Communication Plan Execution

A successful Stakeholder Communication Plan should be executed with clear, two-way communication as a priority. Each stakeholder may require various communication methods that account for accessibility issues; for instance, not every stakeholder may have robust access to internet, so barriers to communication must be accounted for. Ensuring that stakeholders can communicate back, ask questions, suggest feedback, and give testimony to the success of your project or data set is vital to successful communication.

3.4.1 Performance

The performance of stakeholder communication can be clearly evaluated if every method of communication is tied to a metric. Assessing metrics is not always possible, but is a goal for communication. These metrics are extremely useful in measuring which stakeholder audiences were reached and how project funds were spent. Often, tools for communication must be planned explicitly with obtaining metrics in mind. Metrics may differ based on technology or method of communication, but result in quantifiable success. Examples of communication methods and corresponding metrics are listed below:

- 2,500 fliers were mailed to members of the public; the agency received 34 phone calls with follow-up questions.
- A webinar was announced via email to 200 interested professionals. 45 professionals attended live, while 12 people watched the posted recording on YouTube.
- An email announcing the data set was sent to 12,000 interested parties. 9,800 of these emails were successfully delivered, while 1,400 people actively clicked on content within the email (these types of metrics are available from listserv software).
- An Instagram campaign reached an audience of 3,290 people and 90 people reposted the content.
- Within 4 weeks of posting content to a website, 34 engineers contacted the agency with follow-up questions about future projects.

3.4.2 Continuous Improvement

Communication plans should be flexible and dynamic and should allow for continuous improvement and polishing. As stated previously, it is important to focus on two-way communication and provide open lines of communication for stakeholders. Note that additional effort may be necessary to engage marginalized communities. Be sure to consider whether these communities are at the proverbial table and if the planned outreach methods actually reach these community members so they are able to provide information and receive feedback. In some cases, trust must be built through a series of meetings or other methods before meaningful community engagement can occur.

- There should always be an email or phone number to hear questions and feedback from stakeholders. An email alias, such as “communication_feedback@agency.gov”, is one way that allows multiple agency employees to receive and respond to feedback.
- Conduct surveys to understand how well stakeholders are understanding the material. The success of surveys may depend on who the stakeholders are and the level of community trust.
- Listen to suggestions for improvement and assess where improvements can be made. Often a small change can improve communication.
- It may be useful to use a ticketing system to track stakeholder feedback. This may be a help desk or another

ticketing system that allows stakeholder feedback to be tracked, triaged, and responded to appropriately.

3.4.3 Example Communication Plans

Many agencies have formed excellent communications plans that can be used as a template. One example is the Reilly Tar and Chemical Superfund Site Community Involvement Plan. Another example is provided in the Community Involvement Plan, Identifying Priority Restoration Sites for Resilience, Point Hope, Alaska (EA 2021). Use of this community involvement plan is discussed in Section 3 of the Improving Coastal Resiliency in Point Hope, Alaska Case Study.

Every communication approach is unique, especially when engaging with Native American Tribes. Some great examples of tribal consultation policies and protocols developed by various California state agencies can be found on the California Environmental Protection Agency's California Native American Tribal Relations | CalEPA website. Please see the Traditional Ecological Knowledge fact sheets for more information.

4 IMPACT OF COMMUNICATION PLAN ON DATABASE DESIGN AND SERVICES

Database design for an organization's public communications includes some unique criteria. Factors to consider include:

- the purpose of the data sharing project (user/audience engagement plan)
- existing data infrastructure
- technical, personnel, and financial resources
- time constraints
- desired outcome(s) from the system, including
    - access to data
    - information
    - risk warning

- alignment with FAIR principles
- transparency of rules that control which data are public and which data are withheld
- alignment with data accessibility requirements, which will vary depending on the purpose of the database and by agency
- publicly available documentation, including
    - business rules
    - data loading guidelines
    - workflows

Developing a database is complex and requires a good deal of expertise. Additionally, the technology behind databases is rapidly evolving in the commercial and open-source realms. We leave it up to you to choose a platform consistent with your organization's expertise and resources. For more information, please see the EDMS White Paper and the Data Governance Fact Sheet. Discussed below are criteria to consider when designing a database for public data sharing and communication.

4.1 Summary of Database Components

Publicly shared data have two distinct components:

- facts (the most granular form of data)
- interpretations and calculations based on the facts

"Facts" are measurements like particles in air or the toxic bacteria count in a lake while "interpretations and calculations" could be a hazard or risk assessment for exposure to air or water. Examples of implementing both components are an air quality index classification or a safe/not safe hazard indication of whether a water body is safe for swimming.

A database stores transactional data such as field measurements from sensors, analytical results, and other field observations. The database is designed to efficiently accept the data in a well-organized manner. Conversely, as the database size grows, it can be relatively inefficient at querying and transforming the data for analysis and sharing. For example, every time a query is run on a transactional database, the query does the work of converting all the result units to a consistent format and comparing analytical values to screening levels. A transactional database may already exist in an organization to support analysis of environmental conditions by agency staff and regulators and by public entities such as consulting firms, businesses, and interested citizens. Depending on the purpose and outcomes of the data-sharing initiative, the organization could implement a data warehouse.

Data warehouses are designed to transform and organize data from many sources into a cohesive platform, which makes querying and analyzing data easier. The data warehouse simplifies complex queries and makes data analysis fast and efficient. The data warehouse can efficiently form the foundation of a diverse set of data delivery pathways. Examples of data warehouses are open data portals or regional-, state-, and federal-level data clearinghouses.

The data warehouse system will organize and process the data from the existing transactional database to support the end-use need, and a web server or other pathway will “push” or “pull” the data. For example, the existing database might store data from air or water quality sensor measurements. Data flowing from the existing database to the data warehouse would compare the measurements to environmental quality standards, then the web server would provide end users with easy-to-understand graphics such as maps and charts to understand if they should change their behavior to stay safe (stay inside for bad air quality, or not consume fish from a river). As an example, the California Water Boards GeoTracker data management system was originally designed as a database to support regulators managing cleanup sites and water quality compliance. Later, a publicly facing visualization interface was built on top.

Developing a “flow chart” of how data should move through a system will allow for discussions about the appropriate tools. These tools or features may allow users to upload, manage, and view the information. Establishing a specific user role defines how a feature would work for specified parties. For more information, please see the EDMS White Paper.

4.2 Identify Data User (Stakeholder) Roles

Developing features for your public sharing data platform begins with understanding how you expect users/audience to interact with the system. Brainstorming “user stories” can help you develop design specifications for your system, such as data pull scenarios for active user roles or data push scenarios for passive audience members.

First, develop a list of user personas based on the intended goals of your public data system and the role of that user. User personas for a pull system could include environmental scientists who want to download analytical data, a river rafter who wants to know if a river is safe to run based on river height, or a citizen who wants to know if it’s okay to go outside based on air quality. Push users might be drivers who are intending to drive on a road but see an electronic sign indicating the road is closed due to a chemical spill, or members of a town who see news announcements and emails alerting them that their water is unsafe to drink.

A generalized template for developing user stories looks like this: “As [a user persona], I want [to perform this action] so that [I can accomplish this goal].” More information about user stories is explained on the Visual Paradigm website.

Developing system features from user stories can be complex because one story may interactively affect several components of your system. It helps keep the user story as specific as possible.

4.3 Data Management System Roles

Clearly defined data management system roles and responsibilities are critical for an organization’s success in public data sharing. The roles may be defined in regulations to ensure compliance, but there should also be additional guidance documents for further clarification. Typical database system roles are listed and described in Table 1. The goal is to have all parties fulfilling their specified roles so data can still be available to users. If a party is not completing their tasks, another party should act so the data flow is retained. More information about database system roles is presented in the Data Access, Sharing, and Security Fact Sheet.

Table 1. Typical database system roles

Role	Responsibility	Description

Administrator	Oversees the entire system	Oversees the entire system including security, standard operating procedure documentation, database security including backups, ensuring business rules are followed, etc.
Data architect/ developer	Continuously improves features	Maintains the database, bug fixing, develops and applies schema updates, assists the administrator with security.
Data manager	Curates the data	Ensures the data serves it intended goals. Performs quality assurance of the data, ensures pathways are working as intended.
Help desk	Liaison to internal and external users	Provides support for both internal and public data users.
Internal user	Data analyst	Uses the data to develop data reports and interpretations. Can use customized reports or exports to acquire data from the database.
Internal user	Data entry/review	User can provide minor modifications to the system. Should have access to specific tools for managing data entered from other staff or from the public.
Public user	Data analyst	Data consumer who either uses the data for analysis and interpretation, passively uses the data to understand health risks, or becomes aware of the data through push services.
Public user	Data entry	External users that submit data to the platform. This can include regulated communities or the general public. Submission of data can be required by regulation or voluntary.
Audience	None	Receiver of “pushed” or “pulled” data.

4.4 Data Input Quality Control Procedures

Depending on an EDMS’s requirements and accessibility, those users uploading data may be an experienced agency staff member, an external data professional, or a member of the public. Maintaining high quality data in the transactional database requires a mix of engineering design and institutional procedures spelling out specific procedures. Institutional procedures include documenting data loading procedures, training, and a quality control feedback loop performed by a third party within the organization. Engineering design includes data loading applications designed to review data for correct data types, valid values, and input format. Valid values are a defined set of permissible values that may be entered into a given field in a given database. The database will only accept the value as it is defined in a database lookup table. Valid value examples include units and chemical names. Detailed information about valid values is presented in the Valid Values Fact Sheet. Additionally, engineered controls should include automated quality control procedures and queries that summarize and report bad data in the database.

In an example scenario, a data uploader may first check the data types, valid values, and data completeness in a data checking utility. Once the data checker shows that the data comply, the data loader would submit the data upload file for approval by a gatekeeper in your organization who either approves or rejects the data load. Note that the gatekeeper or other staff that manage the database should provide valid value codes to the data loader.

While this approach entails a high level of human interaction, the database system design requirements are minimized. Depending on the complexity and volume of the data, a more sophisticated system could accept the data automatically into a “staging” database where the system automatically reviews and approves the data. In the second case a trained member of the organization may review a report summarizing the checker’s findings. Data input procedures need to align with your data schema, including enforced valid values, data types, parent-child relationships, data quality, and metadata. No matter how your system is designed, you will always be fighting entropy to keep the data organized and usable. For further information on considerations for data quality, please see the Data Quality fact sheets.

Data loading business rules should be well-documented and publicly available. The rules should include:

- definitions of data loader roles and the types of data they can upload, with typical roles including:
    - internal data loaders
    - external data loaders

        - citizen scientists

- training requirements and standards for data loaders
- required documentation for the data loading process, such as:
    - business rules
    - data submittal formats, data types, and valid values
    - steps needed to load data, including quality control procedures/tools

All business rules and guidance need to be written in plain English and be easily findable online. Examples of public environmental data systems with well-defined data loading programs include:

- Washington Department of Ecology Environmental Information Management has a strict process for uploading data to the database, with staff checking implemented for data submittals prior to acceptance.
- IQAir includes a mix of public and citizen science data. This system includes a process to vet contributed data.
- California Water Boards GeoTracker includes system checks of data format prior to submittal with regulator checks used to evaluate data accuracy.

4.5 Database Structures

Understanding the purpose of a database is critical during the development process. The idealized pathway from data collection to communication/posting should be established first. It provides the foundation of how users should manage data. The core database containing measurements should ideally be scalable to allow for growth and schema extensions that incorporate different data requirements. Tools and other features supporting data/information delivery pathways should be designed separately from the database. These will likely constantly evolve as technology evolves and as pathway requirements change or expand. Ideally, plan for your database platform to be flexible in allowing changes as intent of use will continue to change overtime. Establish a database management team to continuously manage and improve the database platform and to answer questions and provide guidance to users and uploaders.

Public communication databases should be developed to:

- be consistent with pathway requirements in a way that meets user/audience needs consistent with user stories
- enable the ability to control access to the data depending on permissions aligned with a user's role
- flag data as publicly available and document the reason why it is withheld, if necessary
- consider user stories and the requirements to support them
- be consistent with FAIR data principles (GO FAIR 2021)
- consider the best practices described in the Geospatial Data fact sheets

Keep in mind that requirements may change during development. Make sure to include budget and time for iterative development. Keep the initial project within the minimum requirements needed to make the system useful. For more information on data management planning, see the Data Management Planning Fact Sheets.

5 References and Acronyms

The references cited in this fact sheet, and the other ITRC EDM Best Practices fact sheets, are included in one combined list that is available on the ITRC web site. The combined acronyms list is also available on the ITRC web site.