



Building the Texas Water Data Hub from the Ground Up

Sam Hermitte • September 2021

What exactly is a water data hub? What should it be? With no roadmap or guidebook to follow, how do you design a hub? What sort of catalyst is needed to really get the ball rolling and build momentum? In Texas, the answer to the last question is steady effort over a number of years followed by a massive hurricane.

Roughly five years ago, the Mitchell Foundation began gathering a small group of Texas water data stakeholders to discuss opportunities to improve decision-making in the water space by improving access to the data that decisions are based on. Through those discussions, the seed for the Texas Water Data Hub was planted. However, before cultivating a hub, it was critical to conduct a gut check with the Texas water community to determine if such an effort would be beneficial to the state.

Enter the 2018 Connecting Texas Water Data Workshop. This event brought together nearly 90 experts from across the Texas water landscape to identify critical water data needs and discuss what the design of a statewide system, if needed, might look like. Through a full day of break-out sessions and discussion, organizers and participants learned that there is an overwhelming desire to make Texas water data more accessible and better connected. In other words, the Texas water community is interested in making water data more findable, accessible, interoperable, and reusable, or FAIR. We were also reminded that the components of the hydrologic cycle are all interrelated, and the associated data are best understood in relation to one another.

“...the components of the hydrologic cycle are all interrelated, and the associated data are best understood in relation to one another.”

With the interest of the Texas water data community piqued and the need for data to be better connected solidified, discussions on how to realize the vision of a

water data hub continued. The initial group of a handful of water data aficionados convened by the Mitchell Foundation expanded to represent a broader array of interests, formally becoming the Texas Water Data Initiative Advisory Committee.

Then, in the wake of Hurricane Harvey — a storm event that devastated portions of the Texas coast and unleashed record-setting precipitation on the state — the 86th Texas Legislature provided the Texas Water Development Board (TWDB)

with additional resources to support the state’s ability to understand and prepare for flood events. Those resources, which were allocated in 2019, support a wide range of efforts to improve the state’s flood resilience, including

- acquisition of complete lidar coverage for the state;
- development of approximate flood risk maps (Base Level Engineering) for the state by 2024;
- installation of additional flood gages in critical areas;
- expansion of the state’s weather observation network, the TexMesonet;
- improved dissemination of online flood risk information through the Flood Decision Support Toolbox ;
- development of a statewide flood planning process; and
- funding for flood mitigation projects through the newly-created Flood Infrastructure Fund.

Additionally, the legislature provided support for the “development of a centralized web resource for flood-related information and water data,” with the explicit goal of developing “a platform (hub) to begin identifying and connecting existing, publicly available sources of water data across the state.”¹ Support for this component of the TWDB’s efforts demonstrated recognition that flood events touch almost every piece of the hydrologic cycle, from surface water flow rates to the groundwater quality of private wells that can be inundated for days, and from precipitation totals to the topography of the landscape that they fall on. Data on



.....
1 TWDB, 2018, Legislative Appropriations Request Fiscal Years 2020-2021: Texas Water Development Board, Austin, Texas, 213 p.

all these topics, and myriad others, are both interconnected and required to better understand, prepare for, and respond to flood events.

Recognizing the heavy lift required to initiate a statewide data hub and the challenges of developing the platform while also standing up new programs in flood mapping, flood planning, and flood financing, the Mitchell Foundation stepped in and provided additional support to accelerate the launch of the Texas Water Data Hub. Thanks to both state support and philanthropic contributions, the seed of the Texas Water Data Hub took root and began to grow in earnest in 2020.

This brings us back to the question of how one designs a hub. In Texas, our team at the TWDB started by taking stock and reaching out. First, we built on and expanded efforts initiated by the Internet of Water and the Meadows Center for Water and the Environment at Texas State University to inventory Texas water data resources. We know that it's critical to identify as many datasets as possible to have a comprehensive understanding of the landscape where the hub will be built. We also recognize that the Texas water data landscape is not static and defining it will be an ongoing task.

Second, our team asked subject matter experts representing 10 open data platforms about data sharing, governance, and hubs. Our goal was to learn from those who have already tread this path; to understand pitfalls to look out for and identify opportunities to collaborate and build on existing work. Through these conversations, we learned that a successful hub should

- add value through problem-solving,
- include clear standards and governance,
- be built with the future in mind,
- take a phased approach,
- empower users,
- be a community effort,
- account for various levels of data literacy and management, and
- serve the needs of data owners.

Next, we completed 11 user research interviews designed to better understand how water data practitioners search for, find, evaluate, use, produce, and share water data. These interviews, which collectively totaled nearly 21 hours in duration, yielded 2,052 utterances that provided direct insight into how real people use real data in real life. All interviewees work specifically with Texas water data, produce or maintain data themselves, and work with data on a daily basis. Through our conversations, we heard why individuals find work with data compelling, learned

that users don't trust high-level search bars and prefer to navigate groupings or search at the record level, and heard about some of the struggles that water data practitioners face.



"If we can bring all that data, bring it to bear and be able to point it at somebody and laser beam it (to) a particular person to answer a question, that's where it's at. **Why are we here if we're not helping people?**"

(Edgar)

"It's still fairly **time consuming to go through and do this exercise. It's just not very efficient...** and the other challenge is sometimes **you just don't know what's out there.**"

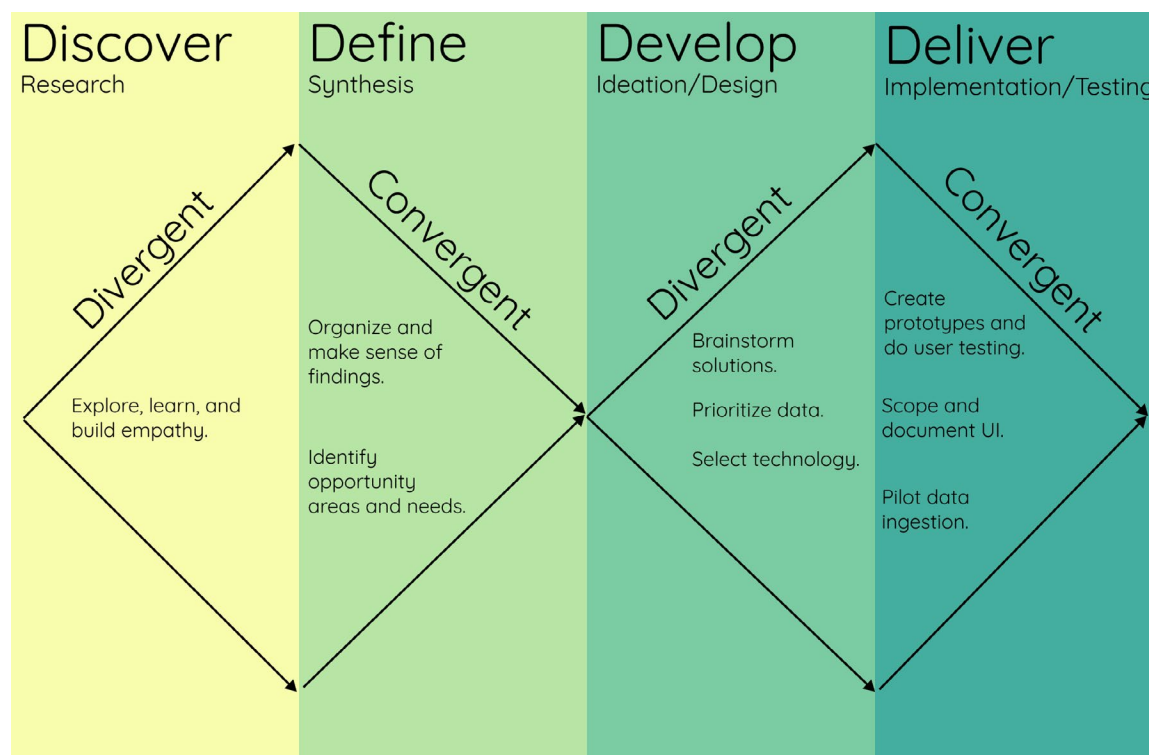
(Erik)



Following these listening and learning sessions, our team transitioned from divergent to convergent thinking. This meant we switched our focus from collecting background research and hearing from a range of perspectives to synthesizing what we heard from practitioners. We looked for patterns, behaviors, and anomalies that we could infer meaning from, using professional judgment to organize ideas and make interpretations that went beyond the obvious. Through this process, we developed our design criteria, consisting of five explicit goals that the Texas Water Data Hub must achieve to be successful:

1. Provide a central location for water data that reflects the entire Texas water landscape
2. Establish automatic and easy ways to share data and updates
3. Provide intuitive methods to efficiently search and download data
4. Emphasize clear communication and documentation to build trust and understanding
5. Assist statewide data interoperability efforts through standards and curated datasets

Collectively, these steps took us through the first half of the design process crafted by our user experience (UX) experts, represented by the first diamond in the graphic below.



With our design principles in hand, we shifted gears again, moving from the convergent synthesis process into a divergent ideation and design phase. For our technical team, this meant spending time researching and evaluating backend technology infrastructure options. Cost and flexibility were two key considerations, as the hub must be manageable from both a technical and financial perspective. Ultimately, we settled on CKAN, an open-source data management system that is used to support numerous other data hubs, including both California's and New Mexico's platforms.

For our data and design team, this meant reaching back out to stakeholders via the Texas Water Data Initiative Advisory Committee to solicit feedback on data prioritization, groupings, and key groundwater and surface water datasets. In preparation for this ideation workshop, we ranked datasets in our inventory using criteria developed for the Texas Disaster Information System, another critical data platform currently in development that will increase Texas' resilience to future disaster events. By using the same criteria, we created consistency across platforms and leveraged the work of partners in the water data space. Once the datasets were ranked, we were able to discuss those that rose to the top with

workshop participants and collectively identify datasets to pilot for ingestion into the hub.

As our team now transitions into the implementation and testing phase of our first year of hub development, we are excited to share our design ideas with users to see how they respond and to begin putting together the backend architecture that will support the hub over the long haul. These are the next steps in a multi-year process to build the Texas Water Data Hub from the ground up, guided by stakeholder input and our mission to create an intuitive system to index, document, search, and access Texas water data. As TWDB Board Member Kathleen Jackson reminds us, “The better the data, the better the science. And the better the science, the better the policy.” We believe that FAIR data are better data, and the Texas Water Data Hub will serve as the cornerstone to a more FAIR Texas water data landscape.

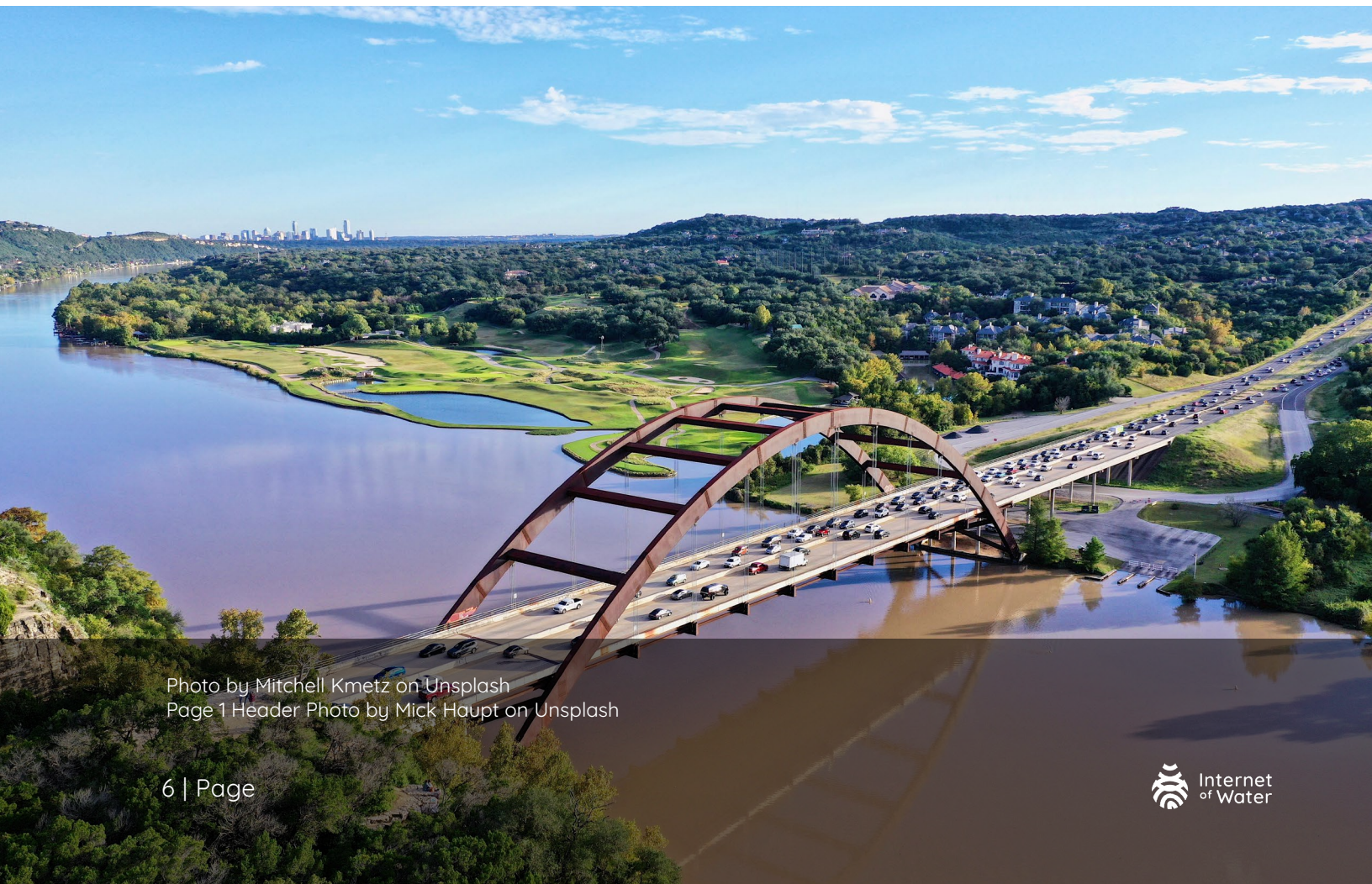


Photo by Mitchell Kmetz on Unsplash
Page 1 Header Photo by Mick Haupt on Unsplash