



**KETOS' submission to the Water Data Prize 2021**

# Table of Contents

1. **Introduction**
2. **Submission 1 - Customer facing Model Website - Communication Category**
  - a. Problems identified
  - b. KETOS' proposed solution
  - c. Impact thesis
3. **Submission 2 - Lead Service Line Mapping Tool - Mapping Category**
  - a. Problems identified
  - b. Proposed KETOS solution
  - c. Impact thesis
4. **Submission 3 - Utility Operator Web Application - Inventory Category**
  - a. Problems identified
  - b. Proposed KETOS solution
  - c. Impact thesis
5. **Alignment with judging criteria**
6. **The KETOS Mission**
7. **KETOS alignment with Water Data Prize goals**

## Introduction

The 2021 Water Data Prize (WDP) highlights key challenges faced by utilities as they prepare to comply with new Federal regulations for Lead and Copper in drinking water. KETOS is submitting its entry in support of utilities' efforts in this regard. We are requesting considerations for judging in the Overall Winner category as well as 3 individual categories - Inventory, Mapping and Communications.

Our submission has been designed with the end users - utility customers and utility operators - as the central focus. While utility customers may intuitively *know* that lead in drinking water is harmful to them, they *may not necessarily have the tools* to self-test or to request testing. Similarly, utility operators may *want* to help their customers with testing and replacement, but they may not *know* which customers are most at risk of having lead pipes. We aim to bridge this crucial gap in utility capability with our submission.

Our [customer-facing website](#) empowers customers to request lead tests from their utility, assess their individual and community lead risk and gain more information about the impact of lead on their health. There are a variety of request types including a mobile water quality testing truck that can be deployed to any community within a municipality which makes the collection of samples equitable and accessible to all communities. This has been of key interest especially to rural utilities that have minimal resources on infrastructure and budgets to invest on analytics.

This website also contains our innovative [Lead Service Line Mapping Tool](#), developed by the KETOS team for this competition as a means to bring increased awareness. The tool has been built from publicly available City of Pittsburgh Lead Service Line data and can be replicated for any other utility with similar data lakes. Here, customers can visualize the public/private lead service line distribution in their home and get more information about lead in their neighborhoods and school districts.

All customer interactions on our customer-facing website flow into our utility-operator [web-app](#). Here, utility operators can monitor incoming requests for testing, schedule lead-pipe inspections and pipe replacements for homes that are confirmed as having lead pipes. This inventory, in turn, can be used to update the Lead Service Line mapping tool in the customer-facing website. The entire flow on the operator-side is part of KETOS' Actinventory model, a uniquely developed solution to enhance and supplement the expertise of utilities allowing them the sophistication to create a lead pipe inventory from scratch using publicly available data. An open-source prototype of the Actinventory model has been made available for general viewing at [this location](#) to help grow its adoption.

As an organization, KETOS' mission is well aligned with the goals of what Water Data Prize intends to achieve for the community. KETOS is committed to empowering utility operators, whose life hasn't been about data science and technology while being the masters of all things water. We are enabling utilities to incorporate these solutions into their daily operations without the need for massive budgets, expensive resources from a hiring standpoint and yet achieving

the customer satisfaction score they hope to deliver for their constituents. Part of this effort is also helping utilities provision our flagship product - KETOS SHIELD - to assist with rapid lead testing for the 1st time in communities identified as vulnerable to lead without the need for long wait times with labs and expensive manual labor. We are also committed to guiding utilities to have a value added offering with their legacy equipment while being interoperable with the solutions driving the future like KETOS SHIELD. Creating a holistic approach to their massive data lakes from decades can yield incredible results when combined with our innovative customer communication and mapping/inventory solutions to improve their overall tech stack.

While we have made every attempt to document the details of our submission, we would be glad to answer any questions you may have as you judge our submission. Please contact us at [ganesh@ketos.co](mailto:ganesh@ketos.co) or [kedar.dabhadkar@ketos.co](mailto:kedar.dabhadkar@ketos.co) with inquiries about our submission.

## **Submission 1 - Customer facing Model Website - Communication Category**

### **Problems Identified:**

After a careful review of the prize guidelines, we identified the following gaps in existing customer-facing websites that aim to communicate Lead Health Risk (LHR):

1. A number of existing websites on LHR (e.g. [Reference](#)) are largely filled with links and text that convey information about LHR but **do not catalyze consumer action** towards solving this problem.
2. Text and infographic heavy websites on LHR also **do not optimize for short-consumer attention spans** (the average website has 6 seconds to retain attention).[\[Reference\]](#).
3. Customers have **different communication needs based on their level of engagement**. In our estimation, existing sites on LHR communication do not do a great job at making content accessible to different categories of users.

### **KETOS' Proposed Solution**

KETOS designed a [customer-facing model website](#) that addresses the gaps above by aiming to achieve the following:

1. **Surface the most actionable information** on LHR first to customers to combat short attention spans and **increase the likelihood of customers actually taking action** on Lead.
2. Customize LHR information to **different categories of users in different formats** while optimizing the verbosity/information ratio.
  - a. Interactive maps that display neighborhood related lead risk (Mapping tool)
  - b. Interactive web-forms that convey individual lead-risk.
  - c. Infographics with optimum verbosity tailored to a pamphlet/leaflet form factor to aid easy printing and distribution by utilities.

3. **Propose Mobile testing using KETOS SHIELD** as a way to increase access to testing, especially to disadvantaged communities. This has the added advantage of increasing the amount of data available to assist with replacement efforts.
4. **Increase customer engagement** through the use of data stories and the ability to share one's personalized results on social media.
5. **Create a seamless interface for data** to flow from the customer-facing website to the utility operators website.
6. **Increase accessibility** by providing multiple options for changing language, font size and color. The website also automatically adjusts form factor for mobile screens.

### **Impact Thesis**

We believe that a customer-facing website with a bias to consumer action, access to testing and increased accessibility will:

1. **Improve customer engagement** by encouraging customers to actively participate in reducing their individual and community LHR.
2. **Increase the amount of actionable data** available to utilities to assist with their Lead Service Line (LSL) inventory creation and management efforts.
3. Communicate the importance of replacing Lead Pipes to end-customers in a way that is **tailored to their attention spans and levels of engagement**.

## **Submission 2 - Lead Service Line Mapping Tool - Mapping Category**

### **Problems Identified**

Most utilities do not have the resources to develop a service line mapping solution. Although several reasonable mapping solutions exist, there are some challenges with these:

1. **Generalizability:** Inability to generalize maps across multiple utilities is the biggest challenge with the existing mapping solutions.
2. **Rigid data infrastructure:** The most widely used dynamic mapping solutions expect the users to convert their data into a format that their tools support.
3. **Difficult to read:** Maps overwhelm viewers with considerable information at once taking away focus from the most critical messages.

### **KETOS' Proposed Solution**

We combined the best features of existing solutions with our proposals to improve mapping. For data storage, we used Excel spreadsheets and a MySQL database. Next, we used Power BI to consolidate data into a data model and develop our visual story. Finally, we used MS Powerpoint to add design components and backgrounds. Here are the features of our solution:

1. **Universal data storage model:** We propose a dimensional data model (star schema), a data structure that allows users to define relationships between various datasets. We can store our data in several formats - Excel spreadsheets, cloud platforms, databases, and many other sources.

2. **Decoupled data storage and visualization layers:** Our tools allows developers to conceptually separate the data model layer from the visualization layer. This allows users to develop a dimensional data model and visualizations separately. Users can refresh data on a schedule multiple times a day to keep the data updated.
3. **Share information in stages:** Our solution shares insights gradually instead of overwhelming the user with data all at once. For example, requesting the users to choose their neighborhood before showing the service line map.
4. **User-friendly design and work-flow:** Animated guides, tutorials, and user-friendly illustrations of public and private service line materials.
5. **Greater engagement with data:** Our solution allows users to download all the data that goes into making any visual.

Sample mapping application has been deployed at [this URL](#). We also made our code publicly available on [GitHub](#) along with a [detailed documentation](#).

### **Impact Thesis**

Any utility can generate their own service line maps using the documented process, proposed dimensional data model and publicly-shared Power BI files. Power BI is free to use until a specific data size limit, beyond which it is as cheap as \$20 a month for a Pro license. Utilities can bring in their data to replace the existing data sources, and the Power BI report updates all visualizations with a click of a button. The ease of use of this solution, our process documentation and the cheap pricing should encourage widespread adoption.

## **Submission 3 - Utility Operator Webapp - Inventory Category**

### **Problems Identified:**

After a careful review of the prize guidelines and the existing literature, we identified the following gaps in Lead Service Line (LSL) management:

1. Utilities (especially smaller ones) **may have no pre-existing LSL inventory** i.e. they might need to create one from scratch.
2. Utilities might face a dilemma in creating LSL and updating inventories - should they **prioritize proactively contacting customers** for inspection and digging, which is more expensive but directly gives them the information needed (location of LSLs) OR should they **rely on customer self-reporting** followed by testing, inspection and replacement.
3. Utilities might not have the **know-how to identify locations likely to have LSLs** prior to first customer contact. This increases the time and budget needed for LSL identification.

### **KETOS' Proposed Solution**

We designed a [Utility-Operator model web application \(webapp\)](#) that addresses the gaps above. We call our approach Active Inventory (Actinventory). It is created by combining KETOS' expertise in an AI approach called Active Learning, the abilities of KETOS SHIELD and existing approaches to Inventory Management in the literature [refs [1](#) and [2](#)]. Our approach aims to achieve the following:

1. Enable Utility Operators to **create and manage an LSL inventory from scratch**.
2. Enable Utilities to **identify likely locations for LSLs using Actinventory**.
3. Propose the use of KETOS SHIELD as a **mobile lead testing device** that can be **deployed at the customer site**. This provides utilities the capacity to **test rapidly and monitor results in real-time** while cutting down testing costs.
4. Creates a workflow that incorporates model outputs, customer self-reports and results from testing, inspection and replacement into **one unified inventory management system**.

### **Impact Thesis**


We believe that the use of our proposed system for Inventory Management will provide utility operators with the following advantages:

1. **Increasing utility efficiency and reducing costs** by decreasing the number of digs needed to locate LSLs. This is achieved by:
  - a. **Machine Learning-based targeted selection** of customer sites with high likelihood of having LSLs.
  - b. Giving utilities the ability to **rapidly test** for Lead at customer sites using KETOS SHIELD, and monitor the results of testing in real time.
2. Creating a **simple, easy-to-understand flow** for Inventory Management that incorporates customer test requests, locations identified by Actinventory and results from inspection and replacement.
3. **Continuous improvement** in the ability of Actinventory to locate LSLs as the inventory is updated using local housing data that is publicly available.

## **Alignment with judging criteria**

Our primary focus has been to propose solutions that we see as having the most real-life impact on end-customers and utility operators. While doing so, we have also endeavored to satisfy as many of the judging criteria as possible. Below, we outline how our proposed solution aims to satisfy each of the criteria.

1. **Health:** In our [customer-facing website](#), we have used a combination of crisp infographics with optimum verbosity, interactive maps and selection buttons to convey the following information:
  - a. Options for testing for lead in water.
  - b. The likelihood of having lead pipes in one's home.
  - c. The impact of lead on the health of Children, Pregnant Women and Adults.
  - d. Homeowner responsibility (vs utility responsibility) in case of LSLs.

- e. Actions that homeowners can take to reduce lead exposure.
  - f. Availability of funding for replacement programs.
2. **Priorities:** Our primary focus has been on improving metrics related to the end-user while achieving the overall goal of finding and replacing LSLs. In the case of our customer-facing website, we have attempted to convey Lead Risk to utility customers while maximizing relevant and actionable information such as test and inspection scheduling and individual health risk assessments. In our mapping tool, we have adopted a locality/neighborhood-level risk mindset. In our utility-operator webapp, we have optimized for ease-of-use within a single streamlined Inventory Management framework and accuracy in identifying potential LSLs using our Actinventory approach.
  3. **Degree of Accessibility:** We have incorporated a number of accessibility options throughout our website. Clicking on the  icon on both websites allows the user to customize their experience accounting for font-size, color preference, displays text and icons in 10 different languages, underlining links and features a day/night mode.
  4. **Level of Engagement:** As mentioned previously - one of our top priorities for the customer-facing website was interactivity and the ability to surface actionable content to maximize customer attention and engagement. When possible (and when consistent with the design) we have used big action buttons with bold font and a call to action (e.g. “Schedule Inspection”, “Check My Home”, “Request Testing in my Area”). We have used multiple types of infographics (static, downloadable) tailored to customers with different levels of engagement. In addition, our Mapping Tool has a number of features (highlighted in the submission) that make it easy and inviting to engage with LSL content.
  5. **Innovation:** We have incorporated a number of innovative design elements in our proposal. The Data Architecture used in our mapping tool has been designed to be general purpose and agnostic to the end-mapping software used. Consequently, our Data model can be plugged into ArcGIS, PowerBI, Basemap, Folium, or any number of GIS/mapping tools. Actinventory combines the concept of Active Remediation ([ref. 1](#)) with the housing data set used in the literature [[ref 2](#)], with one added innovation that should improve efficiency and reduce utility costs - the proposal of the use of KETOS SHIELD as a rapid, mobile tester at customer locations.
  6. **Wowza:** Please see the Innovation section above. We think that’s pretty Wowza! Further to our goal of real-life usability, we have open sourced the data model used in our [mapping tool](#) and the flow-visualization aspect of [Actinventory](#). We would be happy to engage with utilities interested in incorporating these design elements and specific pieces of software into their tech-stack.



7. **Ease of Implementation:** All of our Web-design has been accomplished using standard, easily accessible web-design tools. The Data Architecture in our Mapping Tool does not rely on any proprietary tool. Our prototype visualization was created using Microsoft Power BI, which is free to use until a specific data size limit, beyond which it is as cheap as \$20 a month for a Pro license. Utilities can use our data-model as a turnkey solution by simply plugging-in their data. We have provided instructions on how to do so in our [Mapping Tool documentation](#) along with the [GitHub repository](#). The ease of use of this solution and the cheap pricing should encourage widespread adoption.

Our Actinventory prototype was written and visualized entirely using Python (which is open-source) and other open-source libraries. The approach and the visualization of the [Actinventory model creation](#) process have been shared on a public repository as well.

Broadly speaking, we have endeavored to use open-source tools where possible. In our opinion, none of the tools created, save for the machine-learning aspects of Actinventory and the Data Analysis aspect of the mapping tool, require specialized knowledge. Further, we would be glad to engage with utilities that need help with specific implementation aspects.

## **The KETOS Mission**

KETOS' mission is to create safer, smarter, and sustainable communities around the world where the availability of clean water is no longer a privilege for only a few across the globe.

Our vision is to prevent potential disease outbreaks by leveraging the power of technology and data-driven insights.

Our goal is to create new methodologies where businesses and people transform the way they think about water through smart water management for distribution, safety, and conservation on a global scale.

In service of its mission, KETOS has developed an innovative solution that delivers a vertically integrated water quality and monitoring platform with intelligent hardware, bi-directional connectivity infrastructure, interactive software, and hassle-free services, all for a predictable, flat service fee. Real-time testing and monitoring addresses both water efficiency (leak-detection and usage) and water quality (safety and composition), ultimately increasing overall water availability. With the power of actionable and predictive water intelligence on a global scale, KETOS seeks to solve a number of the world's water challenges with the goal of preserving this quintessential resource for years to come.

The name "KETOS", a derivative of "CETOS" referring to Orca in Greek, symbolizes our endeavor to preserve the ocean's largest mammals, and signifies our commitment to our oceans, our water, and the planet as a whole.

## KETOS Alignment with Water Data Prize 2021 Goals

The 2021 Water Data Prize (WDP) has Lead Water Pipe replacement as the central focus. The WDP aims to provide utilities with tools to easily create, manage and visualize the lead pipe inventory in their service area. It also aims to provide utility end-customers with knowledge about lead and its dangers and resources for investigating potential lead contamination in their homes.

The ability to **rapidly** test water for lead **at customer locations** is likely to be crucial for utilities in their compliance with new federal regulations. In the absence of this capability, utilities will need to rely on **manual testing, slow turnaround times** (on the order of a day/sample), and the **lack of a single digital interface** to maintain test data.

Utilities that do not have a functioning Lead Service Line inventory might face a dilemma in complying with new regulations. Should they focus on testing all customer homes before beginning service line inspections? Or should they focus on inspection, which is more time consuming and significantly more expensive, but addresses their main compliance challenge.

KETOS has solutions that address both concerns.

KETOS' flagship product - the SHIELD, provides utilities with the capability to **rapidly test for lead** (and ~30 other analytes). KETOS SHIELD can be **transported conveniently in trailers to customer locations** and can test customer samples with no time overhead.

KETOS' Active Inventory (Activentory) solution can enable utilities to build and maintain an **inventory of Lead Service Lines (LSLs) from scratch**. This AI-based solution takes the guesswork out of inventory creation and maintenance and allows utilities to focus their time and budget on locations with the highest chance of having LSLs.

Finally, KETOS' innovative software platform can enable utilities to **view customer testing results in real-time**, detect and warn about anomalies and view their LSL inventory through a convenient web-based interface.