# Implementation of the NM Water Data Initiative

**Stacy Timmons**

Associate Director, Hydrogeology Programs

NM Bureau of Geology and Mineral Resources

# Water Data Act (NMSA 1978, § 72-4B)

- Enacted in 2019, from House Bill 651

- NM agencies will **share**, **integrate**, and **manage** our water data
    - Bureau of Geology & Mineral Resources (convener)
    - Interstate Stream Commission
    - Office of State Engineer
    - Environment Department
    - Energy, Minerals & Natural Resources

## Let's make finding water data SIMPLE.

# Growing support and building connections

- Legislative annual funding $110,000 at NM Tech / NMBGMR

- Philanthropic contributions of $25,000 from Healy Foundation

- NM becomes a Pilot with Internet of Water project at Duke University

- **New!** Leveraging state funding for additional $300,000 in federal funds under new US BOR WaterSMART grant (data project in Lower Pecos River region)

**New Mexico Water Data**
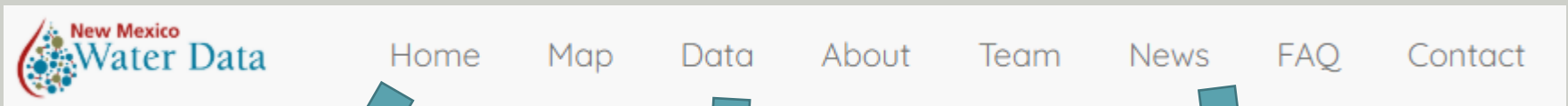
# Recently completed: Plan for Implementation



GOALS

1. Provide IT support to improve data sharing

2. Develop robust IT and data infrastructure plan

3. Refine data standards

4. Continue stakeholder engagement opportunties

# NEWMEXICOWATERDATA.ORG

- Inventory of data and information sources now linked online
- Launched initial Water Data Service – CKAN as our data catalog



Re-construction underway

# Example website: https://www.vvg.org.au/

# General plan for data sharing

1.  Agencies collect, maintain and host data
    –   Only share data when ready / QC'ed
    –   Align with data standards as much as possible

2.  Share data: Work with WDI support team to connect to cloud based CKAN

3.  Integrate data: WDI support team help build transformations

3.  Public access to data: Data available to build tools, analytics, or download and use through web interface

# NM Water Data Pipeline

## Water Data Workshop

**Jake Ross 4/28/20**

# The Problem

```
┌──────────┐   ┌──────────┐   ┌──────────┐
│   CSV    │   │   XML    │   │   API    │
├──────────┤   ├──────────┤   ├──────────┤
│ Meta Data│   │   Meta   │   │ Metadata │
└──────────┘   └──────────┘   └──────────┘
       │              │              │
        ╲             │             ╱
         ╲            ▼            ╱
          ╲     ┌──────────┐     ╱
           ╲───▶│    ?     │◀───╱
                └──────────┘
                     │
        time, depth  │  Metadata
                     ▼
    ┌─────────────────────────────┐
    │         Consumers           │
    └─────────────────────────────┘
```
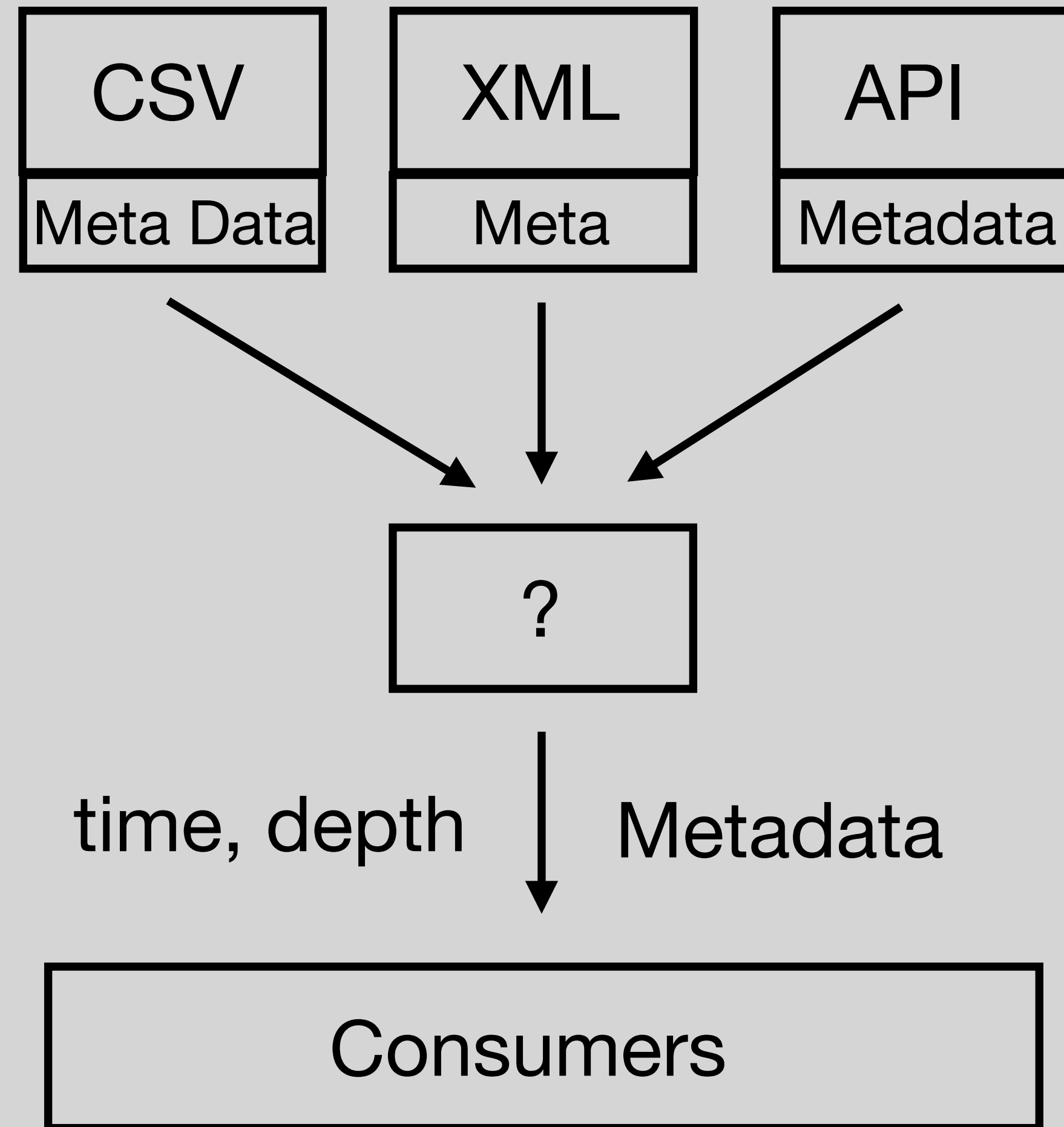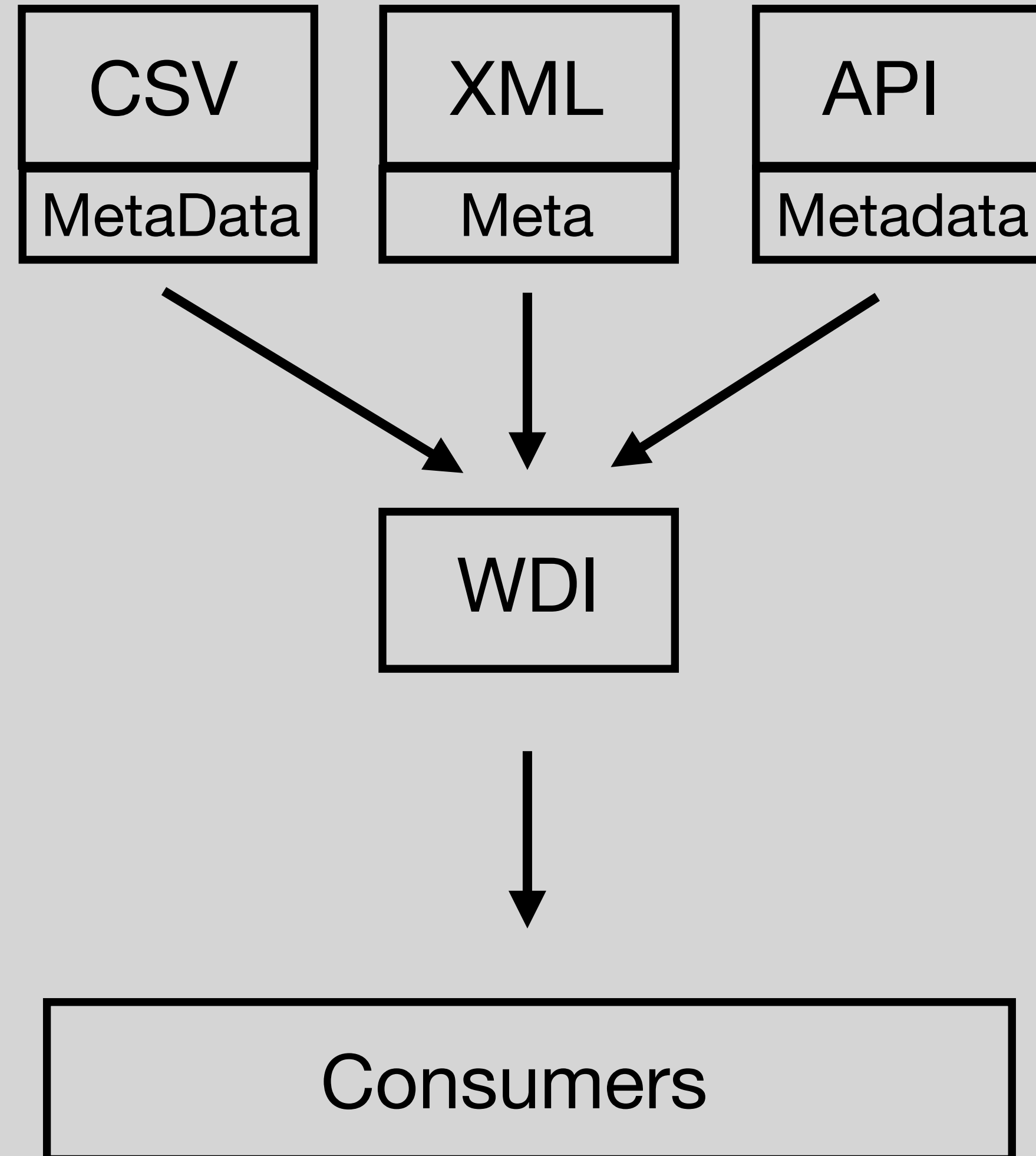
# A Solution

# A Solution
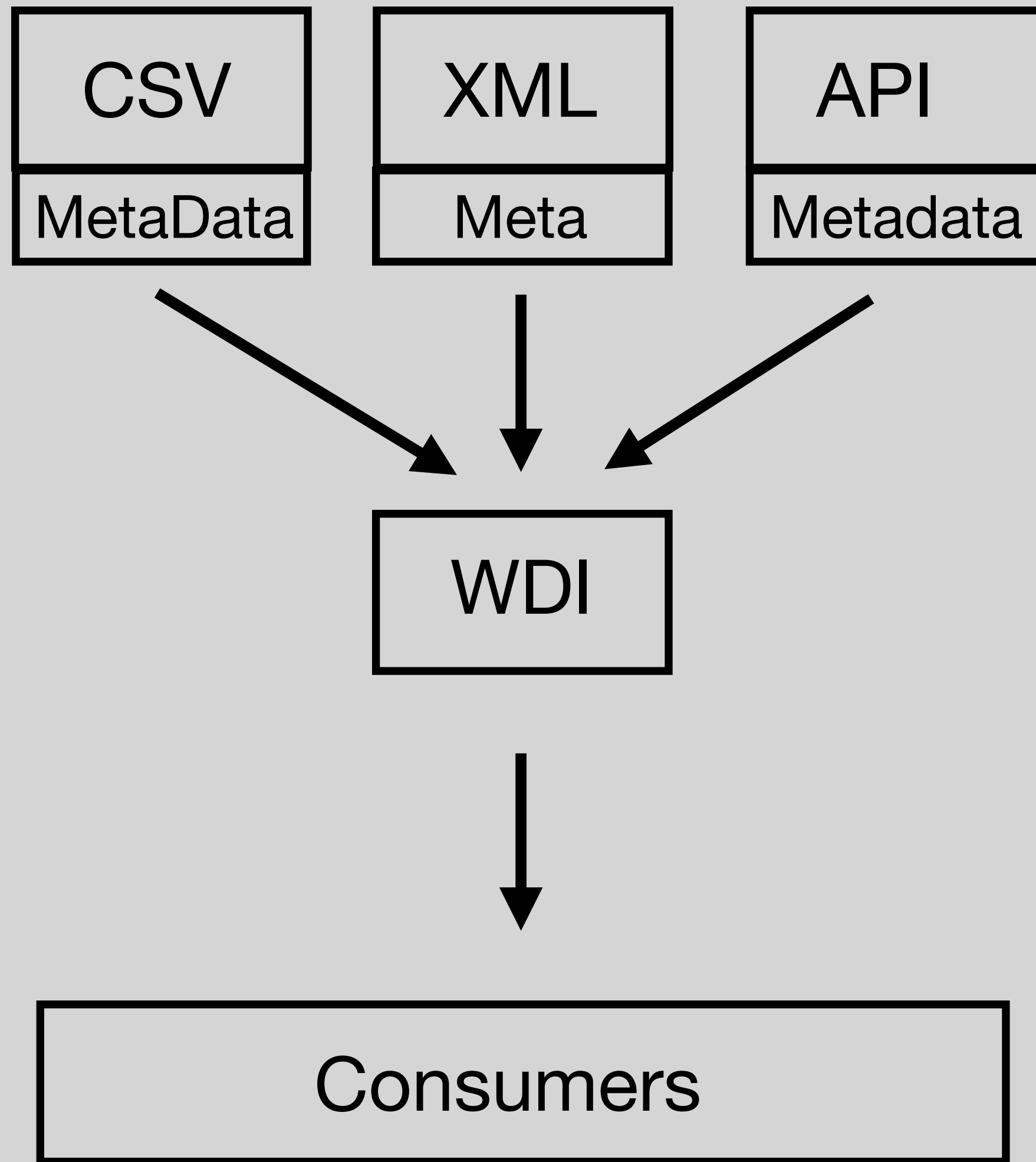
# Who are the "Consumers"

- Applications

  - Web applications, custom desktop apps

- Users

  - Scripts, Browser

- Other APIs

- Integration with Regional and Federal services
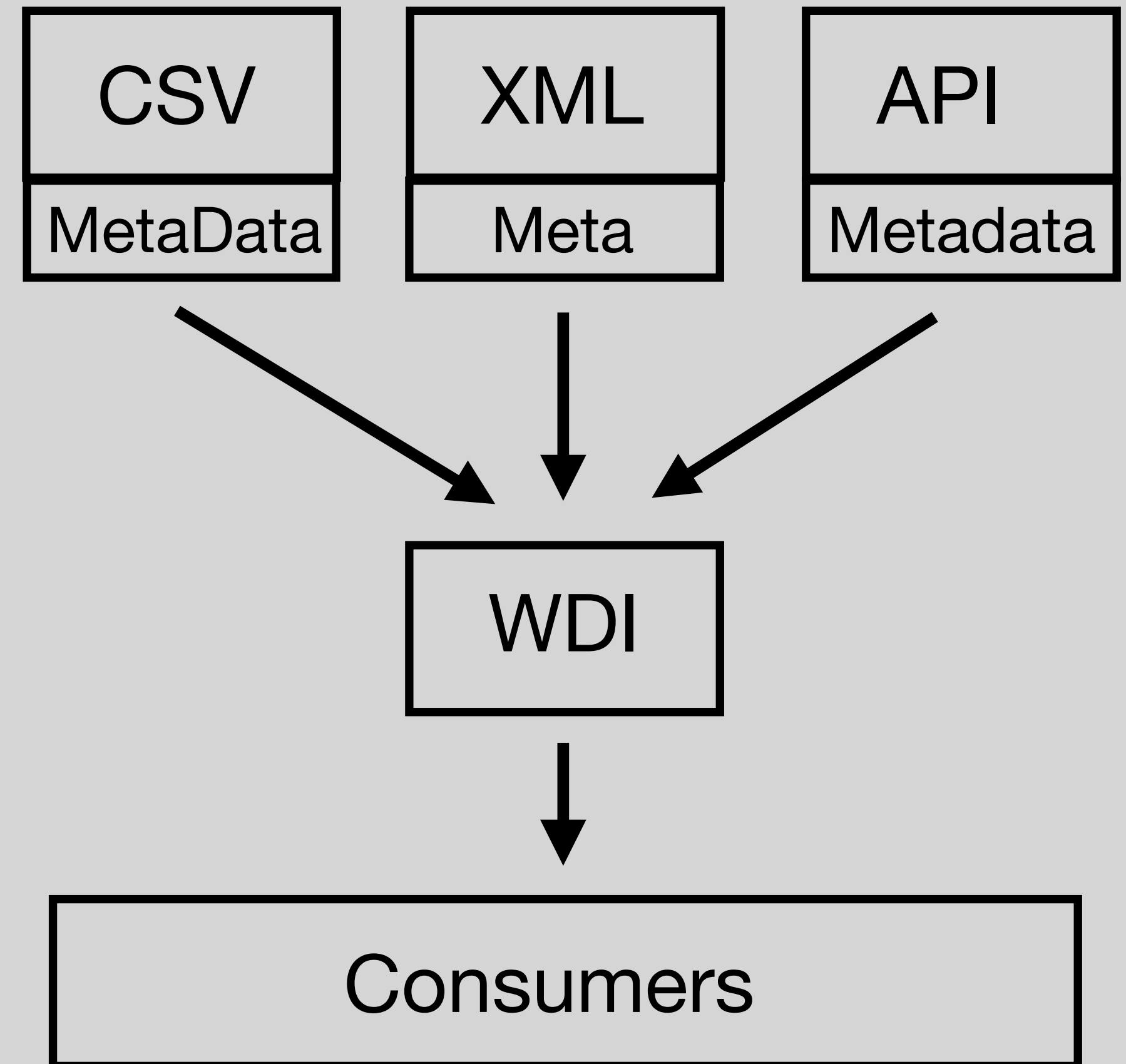
# WDI Process

## Inside the Icebox

- **Initial Review**- compliant? complete? applicable? services? Manual and automated QC tools
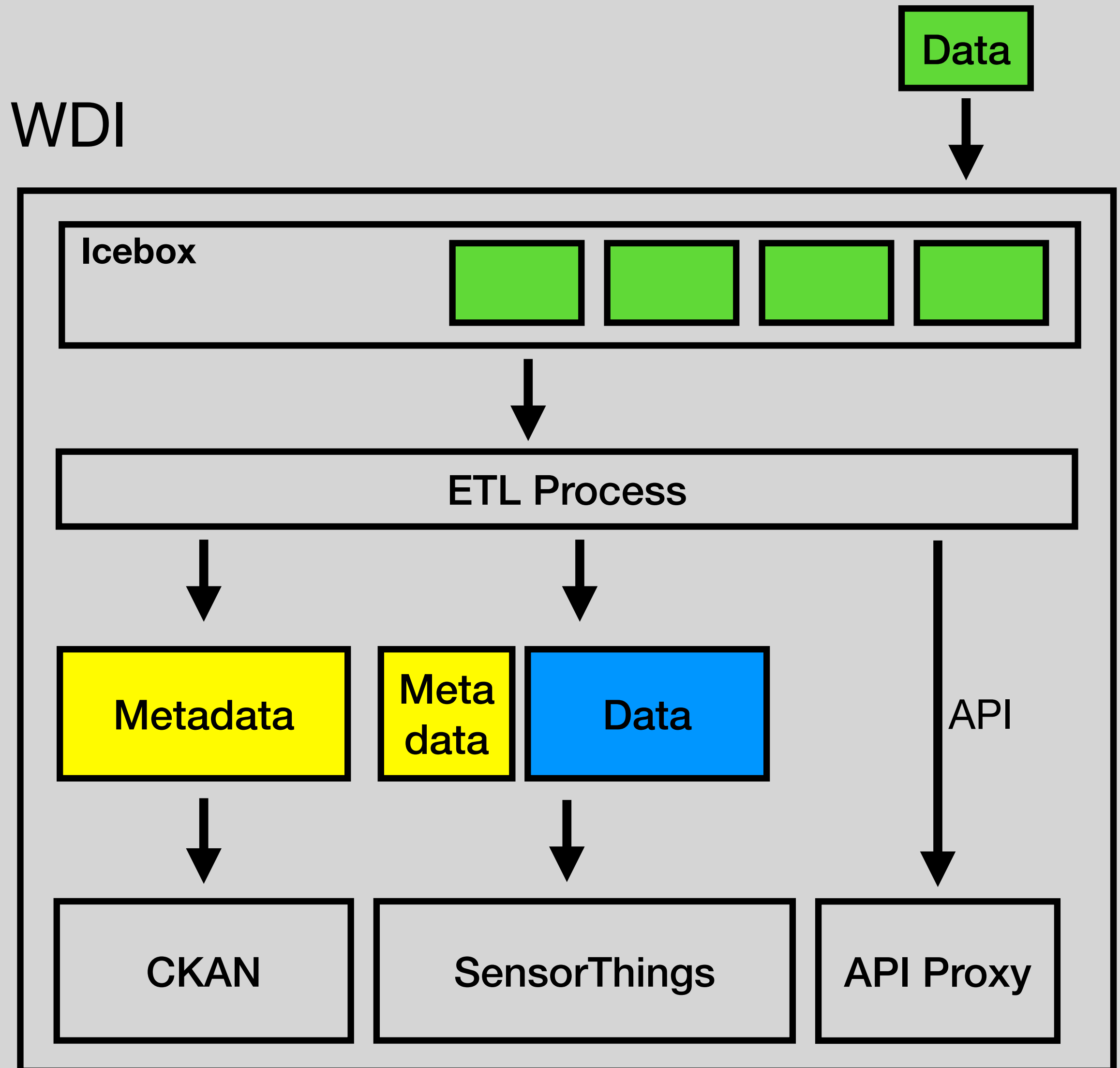
- **ETL**

- **Serve** - Services provide data and metadata to users in an easily consumable standardized format

**Icebox Queue**

**Initial Review**

Metadata **ETL** Built using custom and existing tools e.g. CKANAPI, ChillImport

**CKAN**

CSV, XML, etc

API

**ST**

**API Proxy** translate source API to SensorThings

**Serve**

**Consumers**
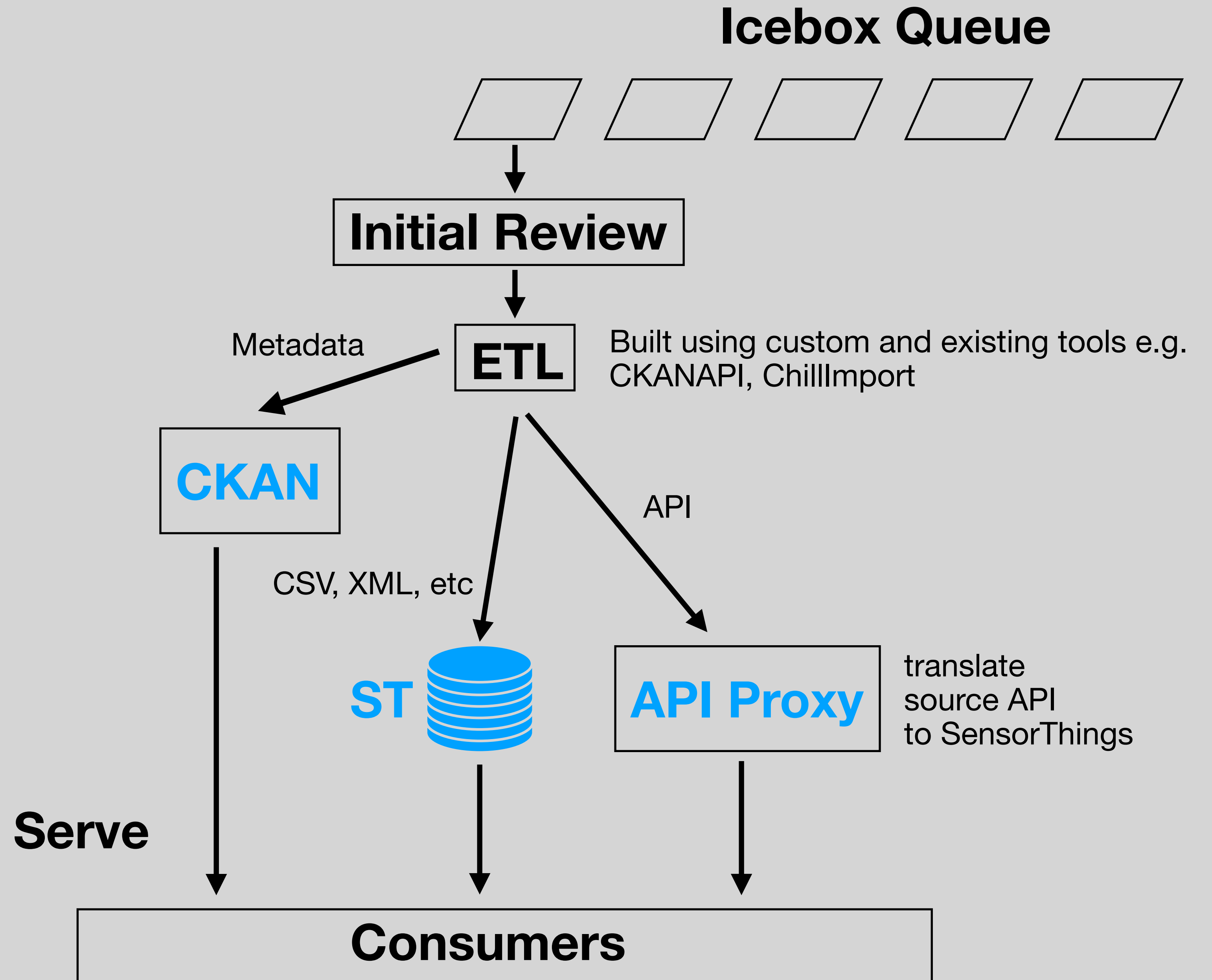
# Phase Two
## Move to a more federated system

- Adapt tools and procedures developed in Phase One

- Each agency has an WDIConnecter

  - used to manage upload/ETL process

  - adds metadata to CKAN

  - adds data to a ST instance

- No need to use API proxying

  - Standardization happens at ETL stage

- Thin WDI management layer
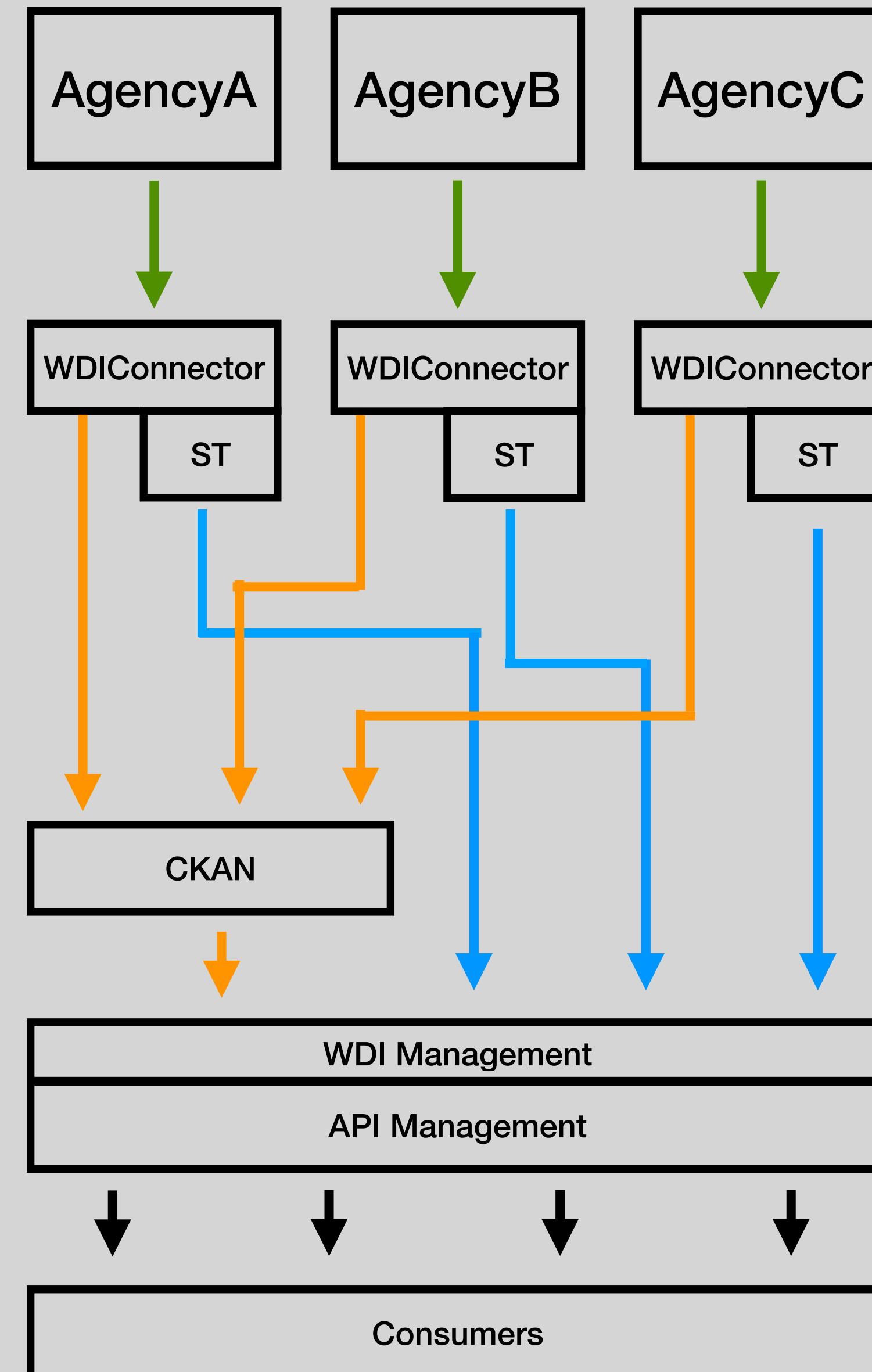
Questions?

# The Problem
## and a solution

- Problem: Share water data and metadata efficiently to the public and researchers

```
┌──────────┐                              ┌──────────┐
│          │                              │          │
│   Data   │ ───────────────────────────▶ │  Users   │
│          │                              │          │
└──────────┘                              └──────────┘
```

- Solution: Common Data Standard + CKAN + SensorThings API

- How will my data be added to WDA?

- What is the process for different data scenarios?

# Terms

- WDA - Water Data Act. This project

- CKAN - Comprehensive Knowledge Archive Network. The data catalog and metadata store for this project

- CKAN "Icebox" - Data staging area. Submitted datasets are stored here before being made public.

- ETL.  Extract, Transform, Load. The process of translating and importing a dataset.

- API - Application Programming Interface, aka simple way for applications to talk to a database via the internet.

- ST API - or STA. SensorThings API, an open source specification for serving location based time series data.

- API Proxy - On the fly translation of an agencies existing API to  SensorThings formatted API.
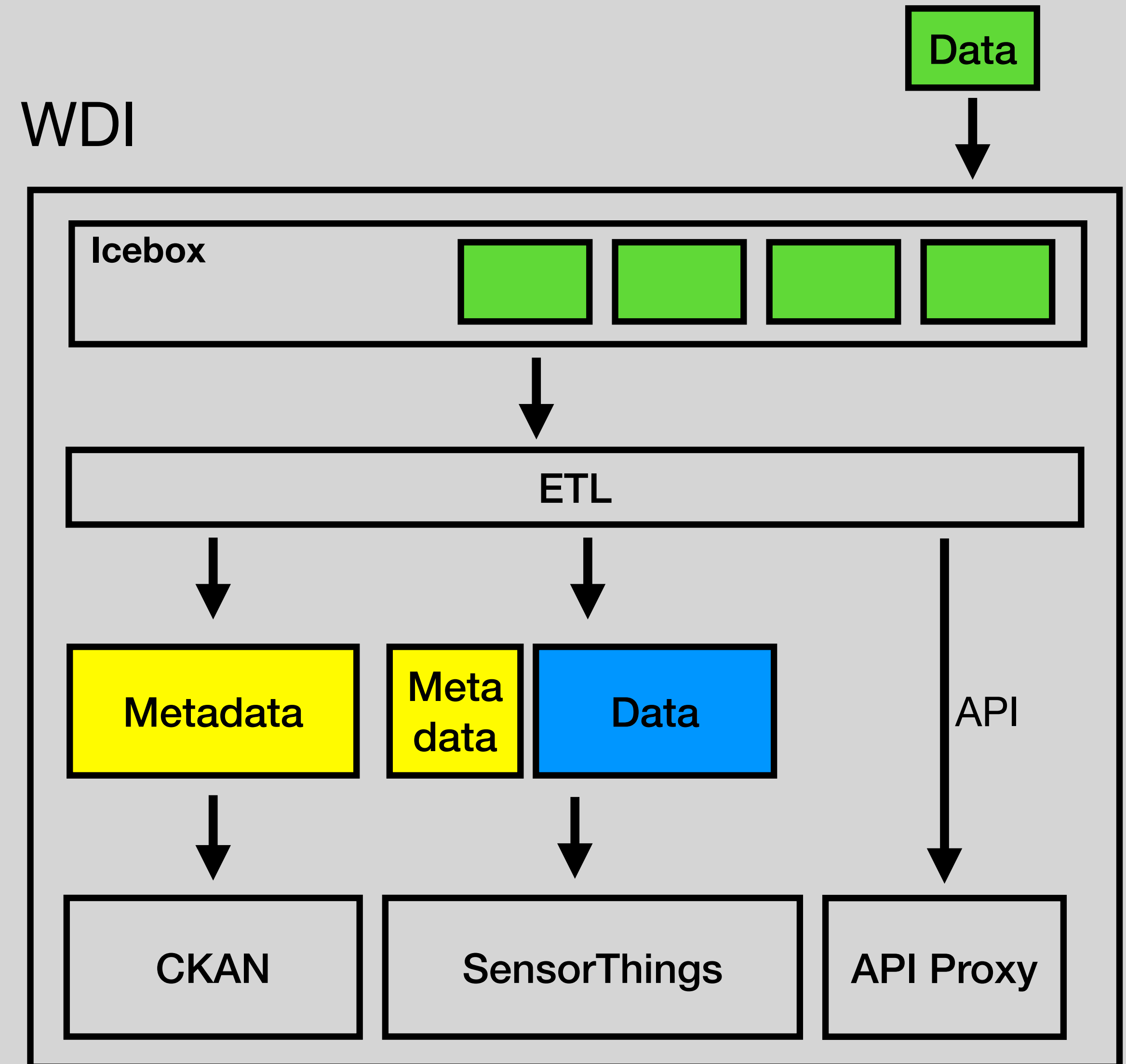
# Your Data
## formats: CSV, XLS, XML, JSON, database dumps, API, etc.

- You have location based time series data

  - that you what to periodically add to WDA

  - or you have a public API

- Data is compliant with the Common Data Standard

- Data (including metadata) is submitted to WDA by a registered user

- Data is now in CKAN Icebox and WDA takes over…

# Data Scenarios

For all scenarios metadata is extracted and added to CKAN

- Text files; csv, xml, txt, etc…

  - Added to SensorThings

- API

  - API Proxied. On-the-fly translation to SensorThings.

  - Served as is. Integration happens at application layer (non-ideal if API is not standardized)

# Potential Data Submission Process

- User registration

  - person or persons responsible for uploading data need to register for user accounts at NM Water Data CKAN website.

- Web form

  - built custom or…

  - built on top of /adapted from CKAN submission tools