

Introduction

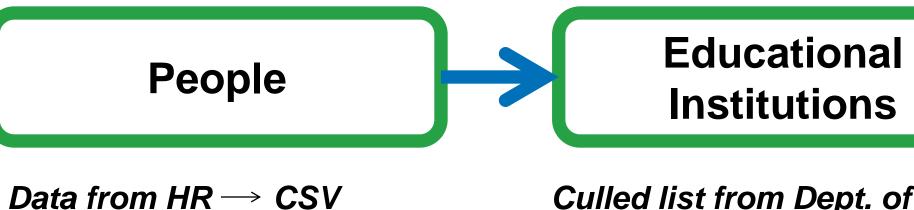
With more than 1,900 federal staff across 13 locations, EPA's Office of Research and Development (ORD) relies on geographically dispersed teams to identify, understand, and solve current and future environmental problems.

We needed a searchable database cataloguing the skills and expertise of ORD federal staff to promote integrated, transdisciplinary research across disciplines, geographic locations, and organizational structure.

We chose VIVO which gives us:

- an expertise location system to help with workforce planning and scientific collaboration
- an option to connect with other systems and pull in preexisting data
- a place for 'elusive' early career data (past employment, training, etc.) and more readily available current data

Data Ingest



↓Manual cleaning ↓Macros for formatting *Python code for parsing data and formatting

Manual Ingest with SPARQL Constructs

Mapping HR fields to VIVO *Creating a new ontology *Editing existing ontology



Education→ CSV Split into: -University, College, & Institute

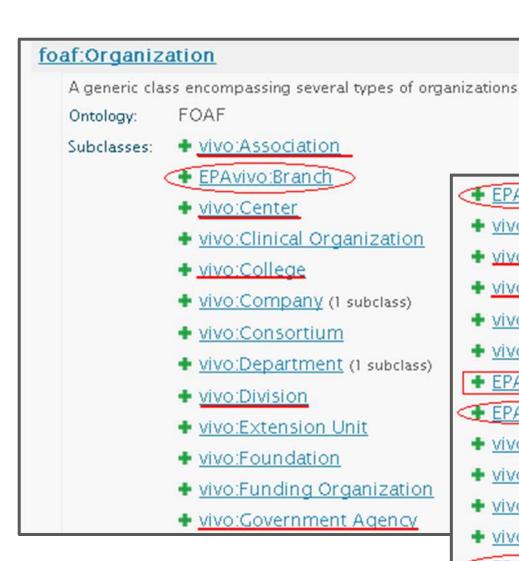
Manual Ingest with SPARQL **Constructs** *Placed in separate graphs *Hidden from browse & index New institutions manually added *Started creating URIs with identifiers

EPA Ontology

The core ontology sets were descriptive and flexible enough to handle most of the data but there were areas where customization was required. This led to the creation of EPAvivo.

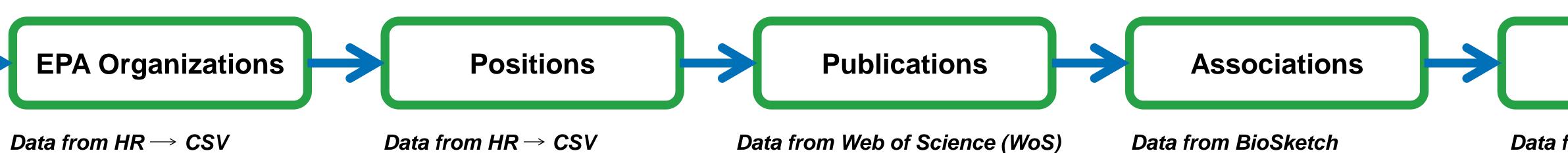
We added data properties for Human Resources (HR), identification, and curation purposes. Examples include: HR ID; Organization ID; Position Curated; and Web of Science ID field, which we use as an alternate identifier for journal articles.

We also customized visibility and editing levels of properties holding HR data. VIVO is not the definitive source for this information, but its presence allows for identification of individuals as well as the aggregation of expertise based on this data.



Moving at the Speed of Government: VIVO Implementation at EPA's Office of Research and Development





*Ontology changes to accommodate ORD Orgs Manual Ingest with SPARQL Construct -Placed into separate graph *First time dealing with

object properties -Made with separate triple file Updated with HR changes

Data from $HR \rightarrow CSV$ ↓ Consolidated list of positions *Pulling individuals Job Title & creating positions upon ingest *New Position Classes in

Ontology RDF file made manually -Placed into separate graph *Date/Time interval ↓ one URI for start date -developed template for triples Users may add previous & concurrent positions -changes stored in kb-2

<u>vivo:Position</u> Particular position in and responsibilities Ontology: VIVO EPA Organizations created Subclasses: 👘 💻 EP EPAvivo:Group to supplement existing + vivo:Hospital ontologies + vivo:Institute 🔸 Vi vivo:Laboratory (1 subclass) 🔸 Vi + vivo:Library **Relevant organizations** = FF + vivo:Museum used from the VIVO EPAvivo:Non-ORD ontology ➡ EPAvivo:Office > + vivo:Program + vivo:Publisher Non-ORD class created for = FF + vivo:Research Organization use when a person creates + vivo:School a Non-EPA position on EPAvivo:Staff their page + vivo:Student Organization EPAvivo:Unit + vivo:University

Cristina Perez, Zachary Painter, Laura Doyle, and David Lyons

Product of the Office of Science Information Management I ORD-Expertise-Team@epa.gov

↓Manual entry from WoS reports led to messy data

*Removed all data

*Created automatic Ingest from WoS files, adding more metadata -built in check for duplicates

-automated creation of date/time interval -URIs created with type identifiers *Assigned multiple ISSNs to Journals accounting for online, print, and

previous versions Expanding process to ingest publications from other sources Data from BioSketch Manual Ingest using SPARQL Construct

*Included websites and screenshots Manually made files for predecessor/successor and

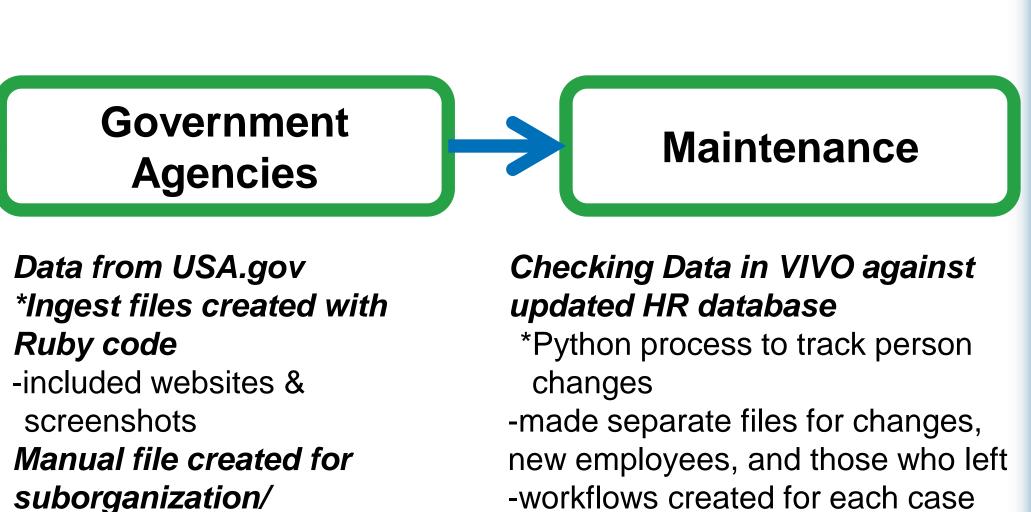
publishing relationships *(linking with journals in VIVO)*

Ruby code relationships

an organization, commonly identified by job title, and normally associat that go with the position. CORE <u>Avivo:EPA Position</u> self inserted/updated position held within EPA	ed with a job descript	tion that details the ta	∳ isks	
itology: EPA Vivo	EPAvivo:EPA	Employee		
o:Faculty Administrative Position	Ontology:	EPA Vivo		
o:Faculty Position	Subclasses:	= EPAvivo:OR	<u>D Employee</u> (1	subclass)
Avivo:Federal Position		Someone who works for EPA's ORD		
position working as a Federal Employee- curated through OMIS/VIVO		Ontology:	EPA Vivo	
itology: EPA Vivo		Subclasses:	= <u>EPAvivo:Fe</u>	deral Employee (7 subclasses)
ro:Librarian Position			A person en	nployed by the U.S. Government
ro:Non-Academic Position			Ontology:	EPA Vivo
Avivo:Non-EPA Position			Subclasses:	EPAvivo:Experts and Consultant
previous position held outside of EPA				+ EPAvivo:Federal Advisory Comm
itology: EPA Vivo				EPAvivo:Permanent
(0:Non-Faculty Academic Position (1 subclass)				• EPAvivo:Public Health Services C
o:Primary Position				EPAvivo:R Appointment
				EPAvivo:Student Workers
				EPAvivo:Term Appointment
L				

EPA Theme

- Customized color scheme:
- 'header-background' images changed to light blue
- 'search-field-and-button' and 'search-interior-pages' images changed to dark blue
- background changed to dark grey (in 'wilma.css'; line 56)
- Added "ORD Research & Expertise" to 'VIVO-logo.png'
- Added a link to the ORD intranet home page at the top right (in 'identity.ftl'; line 38)
- Updated home page text: introductory text on home page (in 'page-home.ftl'; line 46) link to search tips below search (in 'page-home.ftl'; line 63)
- Added FAQs menu page (via Site Admin)
- Implemented Lightweight Directory Access Protocol (LDAP) authentication for user login



- organization within

*Some changes now automated Frequency of check for other process set



EPA Employee: encompasses employees throughout the agency – allows for future expansion of VIVO outside of ORD

Manual updates

Federal Employee: new classes can be added at this level in the future to accommodate nonfederal employees working at EPA (e.g. new class: Contractor)

ORD Employee: subclasses correlate with appointment types used in HR system - these are the classes to which individuals are assigned