Integrated Knowledge and Learning Environment: Harnessing New Technologies to Expand the Diversity of Conservation Participants

Patrick R. Huber (UCD), Allan Hollander (UCD), Matthew Lange (IC-FOODS)

NACCB 2023 – Vancouver, BC
June 27, 2024

prhuber@ucdavis.edu
Introduction – the problem

• Successful conservation planning and implementation involves many stakeholders and data sources.

• Data/information discovery and access, workflow development, and stakeholder identification can be difficult, inefficient, and often duplicative.

• Conservation stakeholders have a wide range of technical skills and backgrounds; how can we help enable equitable access to information and encourage inclusivity?

• How can we integrate data and social networks for improved conservation processes and outcomes?
Introduction – our solution

• Linked data!

• Cyberinfrastructure: machine-readable controlled vocabularies and conceptual structures

• AI tools for assembling and accessing data

• Easy-to-use tools lowering the bar to access
Internet of Conservation (IoC)
Intelligent CI with Computational Learning in the Environment

http://icicle.ai
NSF-Funded AI Institute

Dhabaleswar K. (DK) Panda, PI
The Ohio State University

Follow us
@icicleai
Computer Science
- Software
- AI tools
- Cyberinfrastructure

Domain Science
- Animal ecology
- Digital agriculture
- Smart foodsheds
Internet of Conservation (IoC)
Ontologies and knowledge graphs

- **Ontologies** - formal, machine-readable descriptions of knowledge

- Examples: Gene Ontology, Environment Ontology (ENVO), Sustainable Development Goals Interface Ontology

- **Knowledge graphs** – real world information structured by ontologies
PPOD+

• People, Projects, Organizations, Datasets (PPOD; “Peapod”): who is working on what where and when?

• An ontology describing organizations, their characteristics, and their relationships to other organizations and entities.

• We assembled a knowledge graph of information on organizations that potentially participate in conservation planning and implementation in California.
PPOD+ - ontology

PPOD Core -> PPOD Conservation ->
PPOD California Conservation

https://github.com/PPODschema
As of June 3, 2024
Organizations: 2,742
Programs: 283
Projects: 510
People: 705
Guidelines/Mandates: 152
Infrastructure: 2,276
Datasets: N/A
<table>
<thead>
<tr>
<th>Name</th>
<th>Alias</th>
<th>InfrastructureType</th>
<th>IsPartOf</th>
<th>OrgOwner</th>
<th>OrgManager</th>
<th>OrgUser</th>
<th>IssuesPurpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>180/400-Foot Aquifer Subbasin</td>
<td></td>
<td>Water - Ground Water</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Ground Water, Riparian Water, Ground Water</td>
</tr>
<tr>
<td>Acton Valley Groundwater Basin</td>
<td></td>
<td>Water - Ground Water</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Ground Water, Riparian Water, Ground Water</td>
</tr>
<tr>
<td>ACWD Water System</td>
<td></td>
<td>Water - Water Pipeline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Adobe Creek</td>
<td></td>
<td>Blue - River/Stream</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Adobe Creek (Santa Clara County)</td>
<td></td>
<td>Blue - River/Stream</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Advanced Water Purification Facility</td>
<td>AWPF</td>
<td>Blue - River/Stream</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Agua Hedionda Creek</td>
<td></td>
<td>Blue - River/Stream</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Agua Hedionda Lagoon</td>
<td></td>
<td>Blue - Estuary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Agua Hedionda Lagoon Ecological Reserve</td>
<td></td>
<td>Green - Conservation Area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Ahjumawi Lava Springs State Park</td>
<td></td>
<td>Green - Conservation Area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Aitken Ranch Mitigation Preserve</td>
<td></td>
<td>Blue - River/Stream</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Alabama Hills National Scenic Area</td>
<td></td>
<td>Green - Conservation Area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Alameda Creek</td>
<td></td>
<td>Blue - River/Stream</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Alameda Creek Flood Control Channel</td>
<td></td>
<td>Water - Canal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Alamos Creek</td>
<td></td>
<td>Blue - River/Stream</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Alamo Powerplant</td>
<td></td>
<td>Energy - Powerhouse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Alamo River</td>
<td></td>
<td>Green - Conservation Area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Albany Mudflats Ecological Reserve</td>
<td></td>
<td>Green - Conservation Area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Albany Mudflats State Marine Park</td>
<td></td>
<td>Green - Conservation Area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Albion River</td>
<td></td>
<td>Blue - River/Stream</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Alejandro Canal</td>
<td></td>
<td>Water - Canal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Alexander Subbasin</td>
<td></td>
<td>Water - Ground Water</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Ground Water, Riparian Water, Ground Water</td>
</tr>
<tr>
<td>Alexander Valley Groundwater Basin</td>
<td></td>
<td>Water - Ground Water</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Ground Water, Riparian Water, Ground Water</td>
</tr>
<tr>
<td>Aliso Creek</td>
<td></td>
<td>Blue - River/Stream</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Aliso Creek (Los Angeles County)</td>
<td></td>
<td>Blue - River/Stream</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Aliso Creek Estuary</td>
<td></td>
<td>Blue - Estuary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
<tr>
<td>Alki Sink Ecological Reserve</td>
<td></td>
<td>Green - Conservation Area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Water, Water Infrastructure, Riparian Water</td>
</tr>
</tbody>
</table>

ICICLE
Internet of Conservation (IoC)
Internet of Conservation (IoC)
Interactive Knowledge and Learning Environment (IKLE)

LLM APIs

Knowledge Harvester

FastKG

Pods API

Data DBs

ETL Interface

Workflows API

Visual Studio

Visual Studio Frontend

Visual Studio Backend

Visual Studio

Speech2text

Text2query

Text2vis

KG API (David)

Postgres

Ontologies API (David/Matthew)

PPOD

Ref Arch

Molecules
Interactive Knowledge and Learning Environment (IKLE)
IKLE Visual Studio
IKLE Visual Studio
"Find possible collaborators that work on water quality."
Conversation agent

Large Language Model (LLM)

Ontology/KG
“Find possible collaborators that work on water quality.”
Expanding participation

• AI tools and cyberinfrastructure designed to reusable and made widely available

• Generalized data model allows users to bring their own data and link it to other data

• Voice-activated natural language query tools allow non-data scientists to access information

• Future work: apps for edge devices will make IKLE/ICICLE capabilities widely available
Next steps

• Finalize data federation with four example databases, identify new databases for federation

• Finalize LLM/ontology-based conversation agent

• Continue work on Knowledge Harvest pipeline

• Identify partners for development of new use cases, ensure widespread usefulness and usability
We would like to thank National Science Foundation (#2112606) for ICICLE funding

Contact us: Patrick Huber (prhuber@ucdavis.edu)

Some links:
• Center for Conservation Informatics (https://environment.ucdavis.edu/center-conservation-informatics)
• IC-FOODS (https://www.ic-foods.org/)
• ICICLE (https://icicle.osu.edu/)
• Prototype viz platform (https://vaapifrontend.pods.icicle.tapis.io/#/)