

# New York Climate Change Science Clearinghouse (NYCCSC)

*Extending VIVO ontology and application to facilitate discovery through  
Blacklight and to support climate change science content curation*

Huda Khan, Jon Corson-Rikert, Darcy Branchini, Ingrid Zabel

# Overview

- Background and motivation
- Site demo
- Approach and relevance to VIVO community
- Architecture and extensions overview
- Content and ontology development
- Search and display extensions
- Future work
- Questions

# Background: Addressing climate change

- How will it affect **NY State**? How will it affect **my local town**?
- **Vulnerabilities?** Potential **adaptation strategies**?
- Example: **Flooding** in urban areas due to greater rainfall
- Possible responses
  - Phased withdrawal from floodplains
  - Implementation of federal stormwater management regulations
  - Developing more resilient infrastructure
- End users need to access
  - Climate models and predictions
  - Related policies and documents
  - Relevant data

# What is the NYCCSC?

- Central point of *search and access* to *climate change information* relevant to *NY State*
- Funded by NYSERDA, the New York State Energy Research and Development Authority
- Partners:
  - Northeast States for Coordinated Air Use Management (NESCAUM), Cornell University, SUNY Environmental Sciences and Forestry, NOAA, and the Northeast Regional Climate Center
- Audience to include state & local policy makers, researchers, and members of the public
- Status: Beta release expected by fall
  - <http://climate.library.cornell.edu> (currently still restricted access)

# Site demo

- Primary organization
  - Maps | Data | Documents
  - Context pages
  - Search results and item views
  - Sector pages
- Motivating example
- Faceted search overview
- Geographical location search

# NYCCSC project goals (\*)

- Compile and coordinate ***scientific data and literature*** that will assist ***New York State*** in achieving its ***climate change adaptation*** and ***mitigation goals***
- Provide user-friendly, web-based ***public access to data and literature*** related to climate change science that is relevant to New York State
- Serve as an ***academic climate change data and literature clearinghouse*** that will ***support climate change research activities*** of state agencies, authorities, municipalities, private business and the insurance industry
- Build on existing efforts, leverage significant co-funding, and facilitate the long-term availability of data

# Approach: help people connect the dots

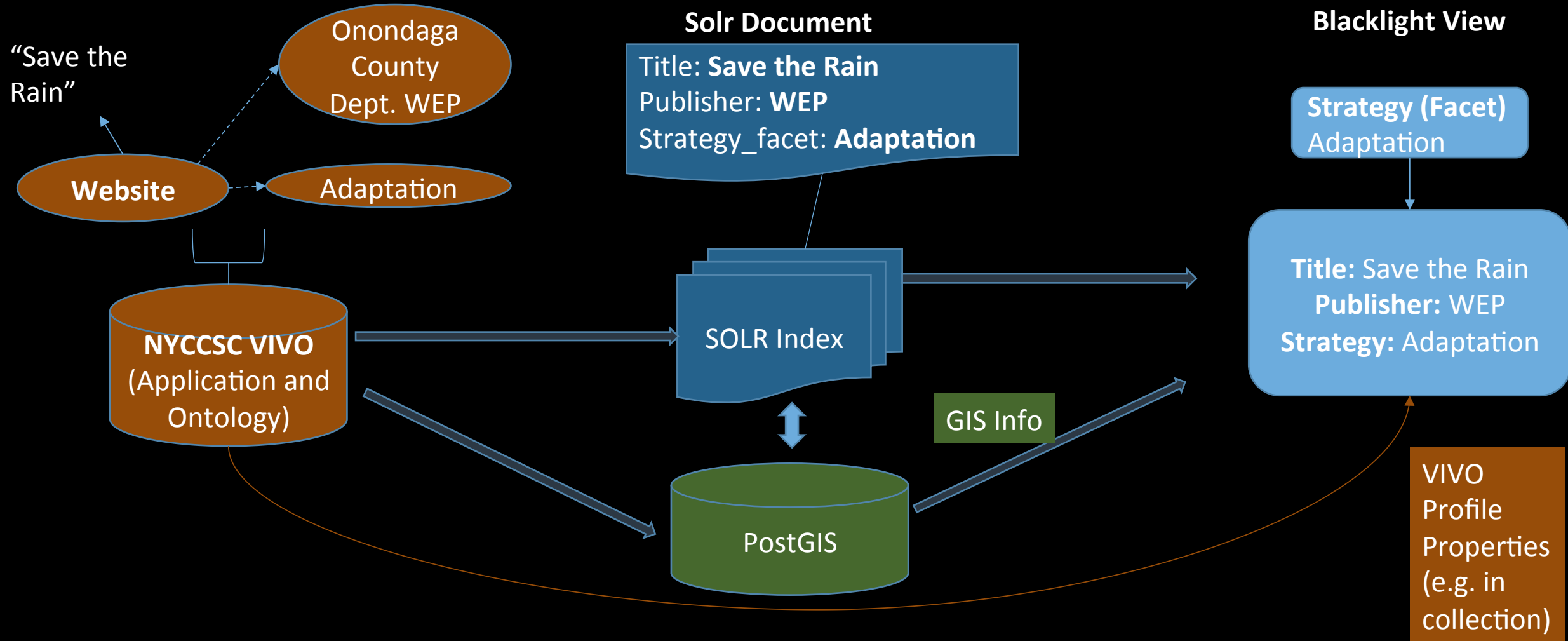
- Audience will range widely in familiarity with scientific concepts and data
- Climate change is a **very broad domain** with profound implications for **regional, state, and local** issues
- **Sector experts** have traditionally communicated well **within their domains** but had few opportunities for knowledge exchange
- Local officials with broad responsibilities for emergency planning and response **need information from multiple sectors and scientific domains**
- Emphasis is on how underlying climate changes influence vulnerabilities and may be responded to via actions that reduce risk

# Relevance to the wider VIVO community

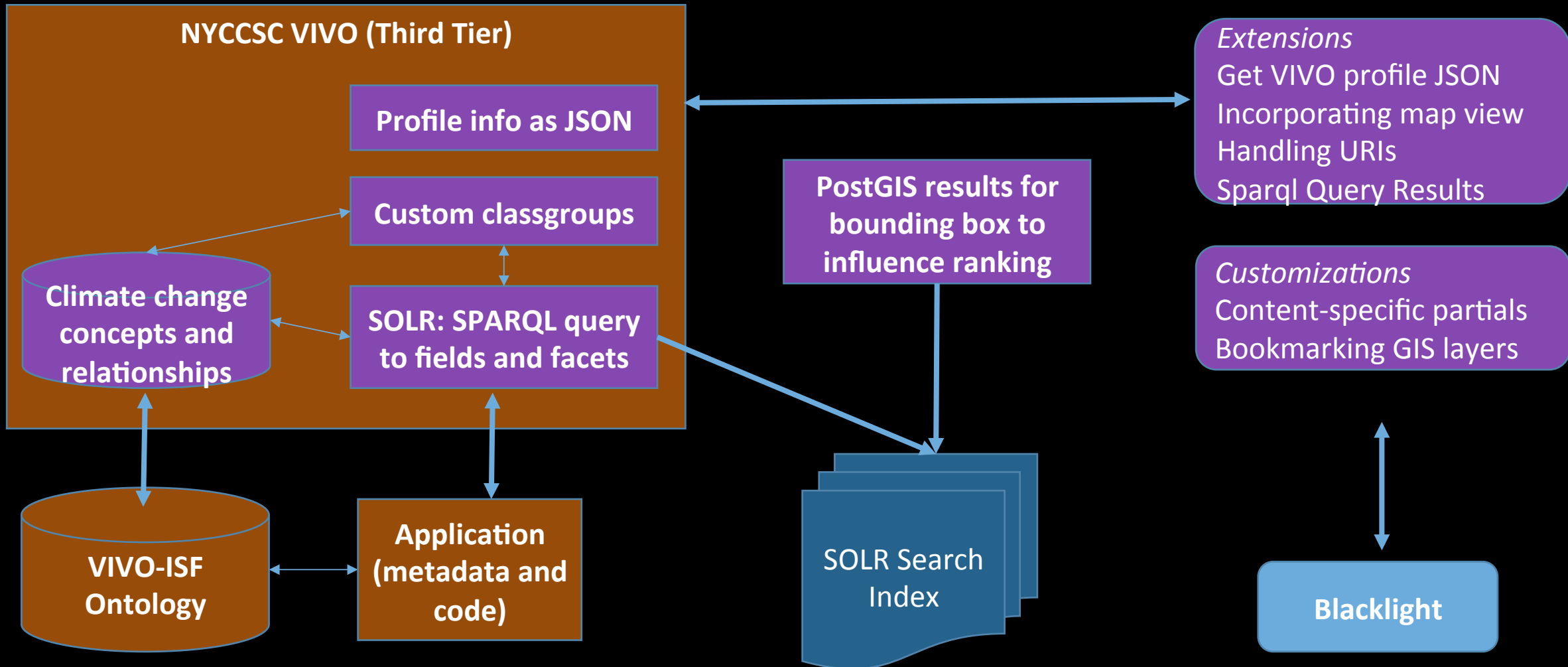
- Adding to VIVO-ISF ontology base by modeling climate change concepts
  - Perhaps useful as a model in other domains of science and/or for integrating data across disparate domains and geographic areas
- Displaying VIVO content via a different front end
  - Exploring the value of semantic relationships to augment faceted search and create richer item-level views
  - Developing new mechanisms for retrieving page-level compilations of data for delivery outside of VIVO
- Customizing VIVO to support spatially-driven search requirements
  - Adapting and populating VIVO's Solr instance
  - Geographical sorting based on PostGIS information



# Architecture Overview



# VIVO Extensions Overview

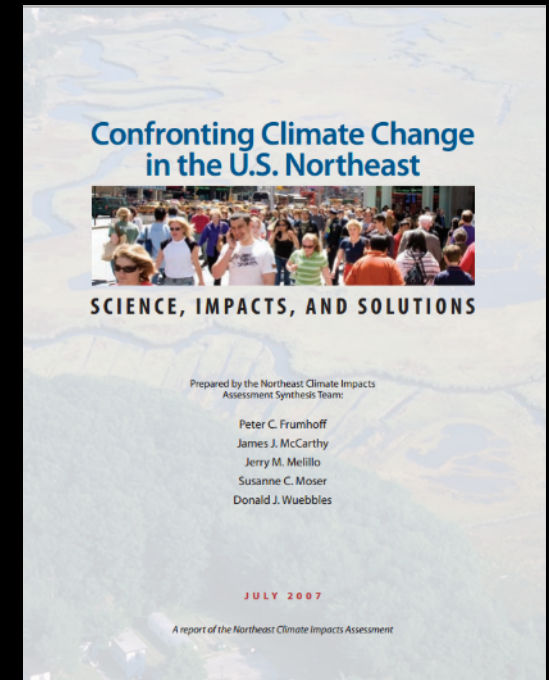
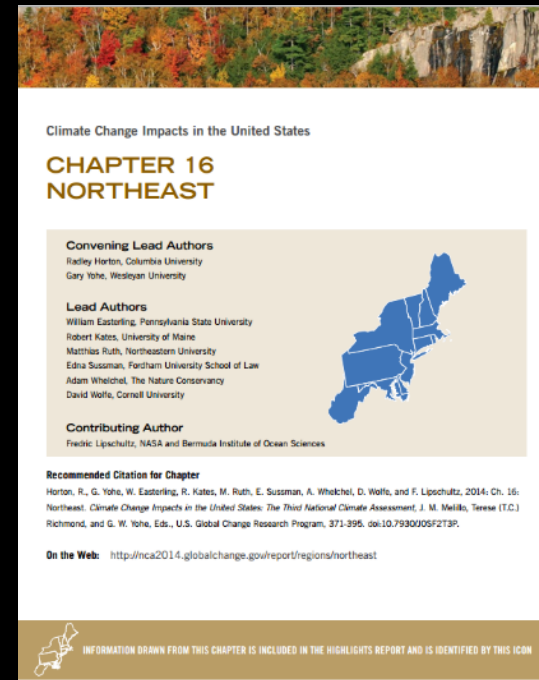
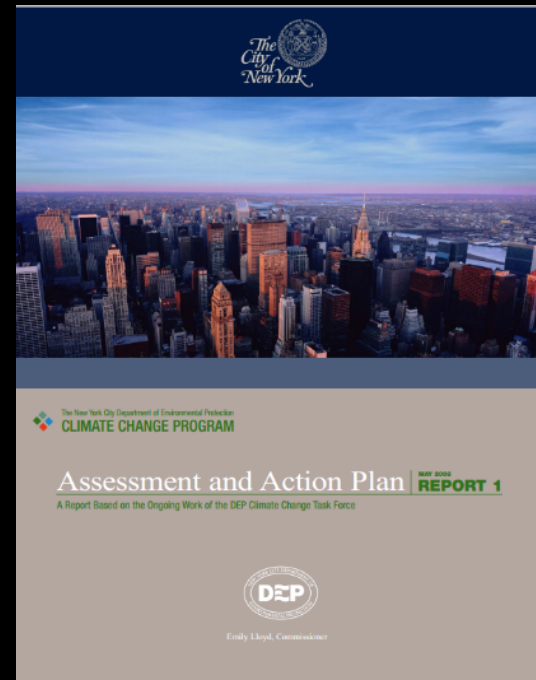
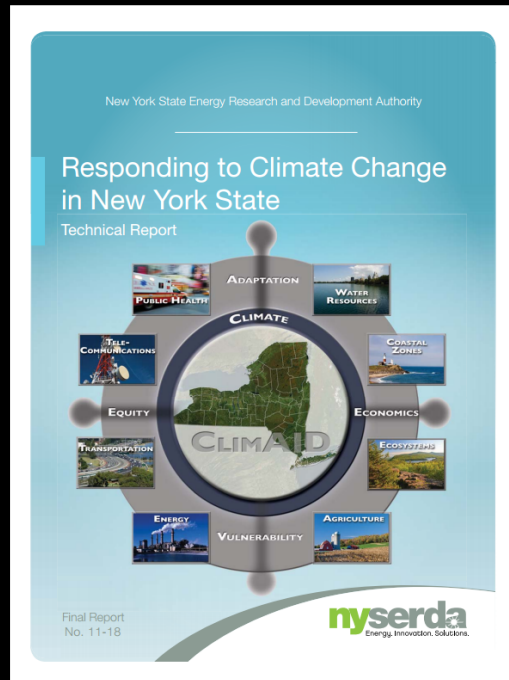


# Content and Ontology development approach

- Guiding factors
  - User-centric: Discoverability, understanding of relationships, context
  - Content-driven: Modified based on available content and the relationships evidenced
  - Curation: Categorizing and querying to evaluate relevant information
- Seeding the Clearinghouse
  - Main content sources helped to generate first version of extensions
  - Concepts and relationships – guiding from awareness to understanding to action
- Examples
  - Classgroup construction to co-function as facets: climate change concepts
  - Granularity: facets work sometimes at the class but other times at the individual level

# Initial Clearinghouse Content

- Four seed documents and their bibliographies



# Initial Clearinghouse Content

- Content relevant to User Scenarios
  - Bridge engineer
  - Leader / elected official for a municipality
  - Public health commissioner
  - Mall developer
  - Coastal Specialist - Office of Planning and Development
  - Graduate student at a university in New York State
- Articles, reports, case studies, maps, data sets, model simulations, planning documents...

# Curatorial interface in VIVO

- Natively provides a rich environment for data curation
- Semantic relationships are richer than simple tagging or linking
  - A strategy ***“is addressed by”*** documents or other resources and ***“implementable through actions”***
  - A climate change ***“influences vulnerability”*** and ***“motivates response actions”***
  - Not just laundry lists of mixed bags of related items
- Need not emphasize the person and organization focus of a typical VIVO
- Through permissions, allows curation of highlights collections and administrative reports
- Limited support for provenance has been added to the data model but not fully automated

Sector	Climate Change	Effect	Strategy	Action
Agriculture	Cloud	Animal or Insect	Adaptation	Built Infrastructure
Buildings	Humidity	Built Systems	Mitigation	Carbon Sequestration or Capture
Coastal zones	Hydrology	Health		Diversification
Ecosystems	Ocean	Natural Systems		Emissions Reduction
Energy	Precipitation	Plan		Energy Efficiency or Conservation
Public Health	Snow and Ice	Specific Opportunities		Forest Management
Telecommunications	Storm	Specific Vulnerabilities		Geoengineering
Transportation	Temperature			Land Use
Water Resources	Wind			Management or Operations
				Natural Systems Renewal
				Policy
				Recycling
				Relocation
				Risk Management
				Social Innovation
				Waste Management

Fundamental organizing elements & flow from the general to the specific

# Ontology development

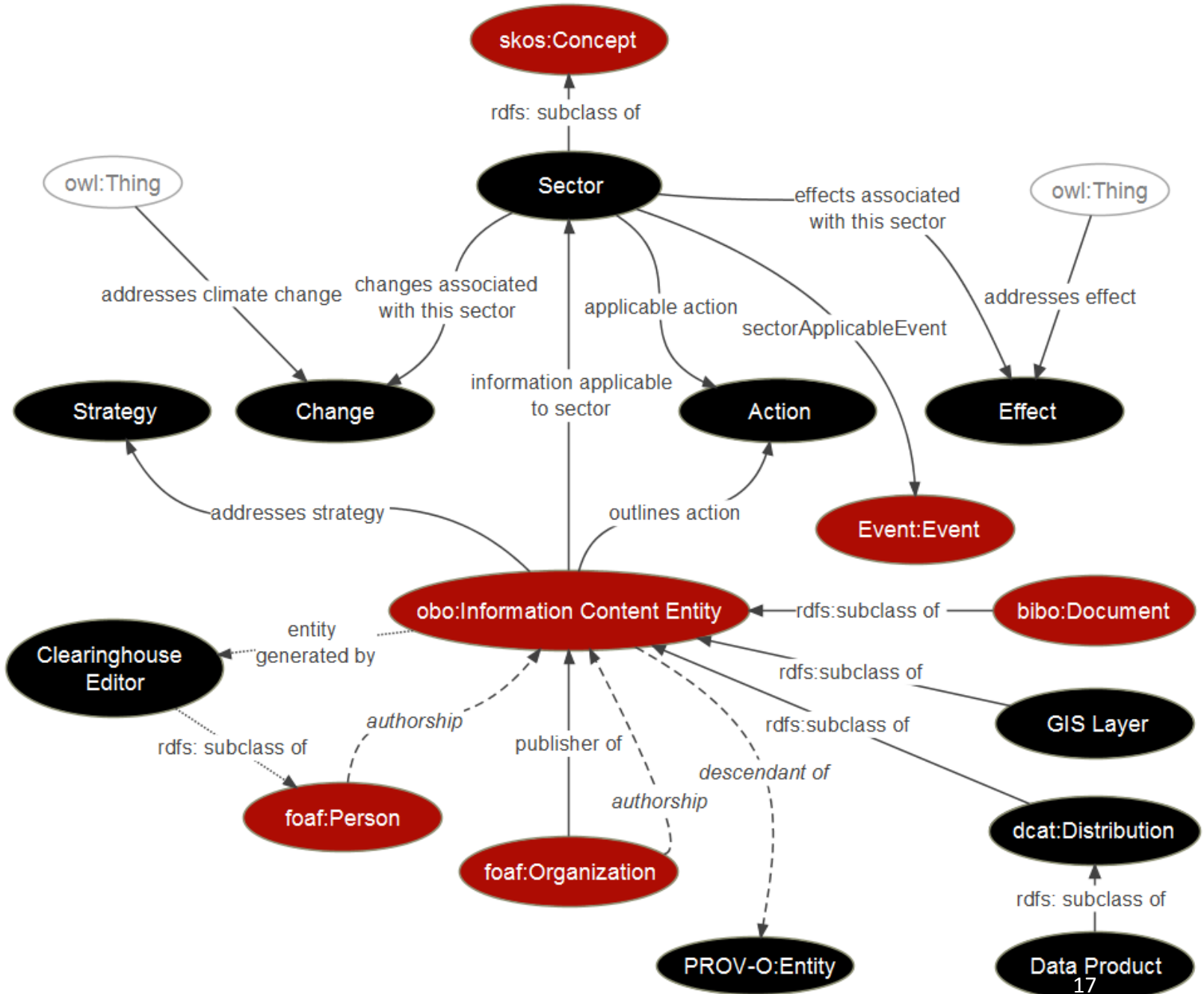
- Granularity of categories
- How would user search/find information
- Curation

Sector	Subsector	Effect	Specific effect	Related climate change
Water resources	Infrastructure	damage to and strain on water treatment infrastructure	New maximum potential stream flow/flooding in large basins	annual precipitation, heavy precipitation events
Water resources	Infrastructure	damage to and strain on water treatment infrastructure	Limitations of urban watersheds to attenuate rainfall inputs	annual precipitation, heavy precipitation events
Water resources	Infrastructure	damage to and strain on water treatment infrastructure	Flooding of wastewater treatment plants	inland flooding, coastal flooding, heavy precipitation events, sea level rise
Water resources	Infrastructure	damage to and strain on water treatment infrastructure	Flooding of coastal water infrastructure, including wastewater treatment plants	sea level rise, storm surge, coastal flooding
Water resources	Infrastructure	damage to and strain on water treatment infrastructure	Stress on stormwater management systems due to increased heavy rainfall events	heavy precipitation events
Water resources	Drinking water supply	changing demand and competition for water	Changing demand	annual temperature

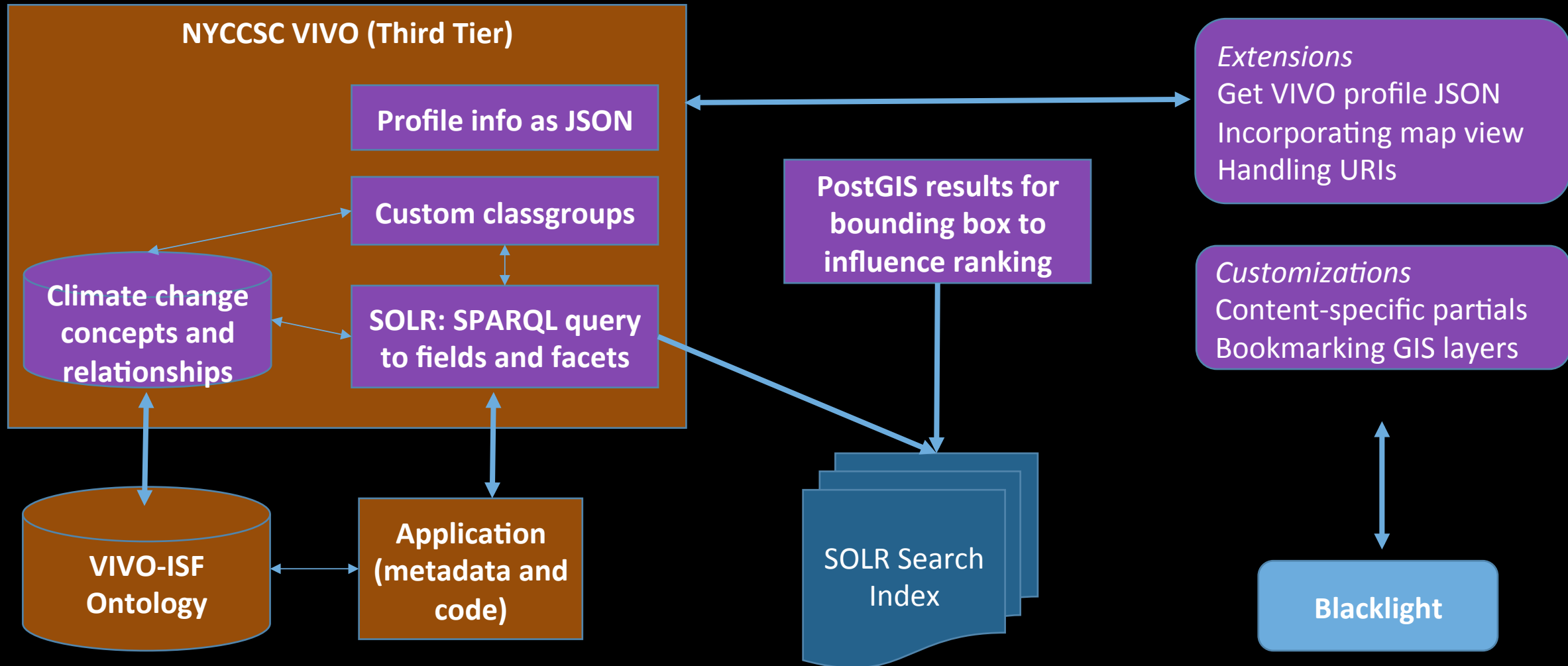


# Simplified Ontology Overview

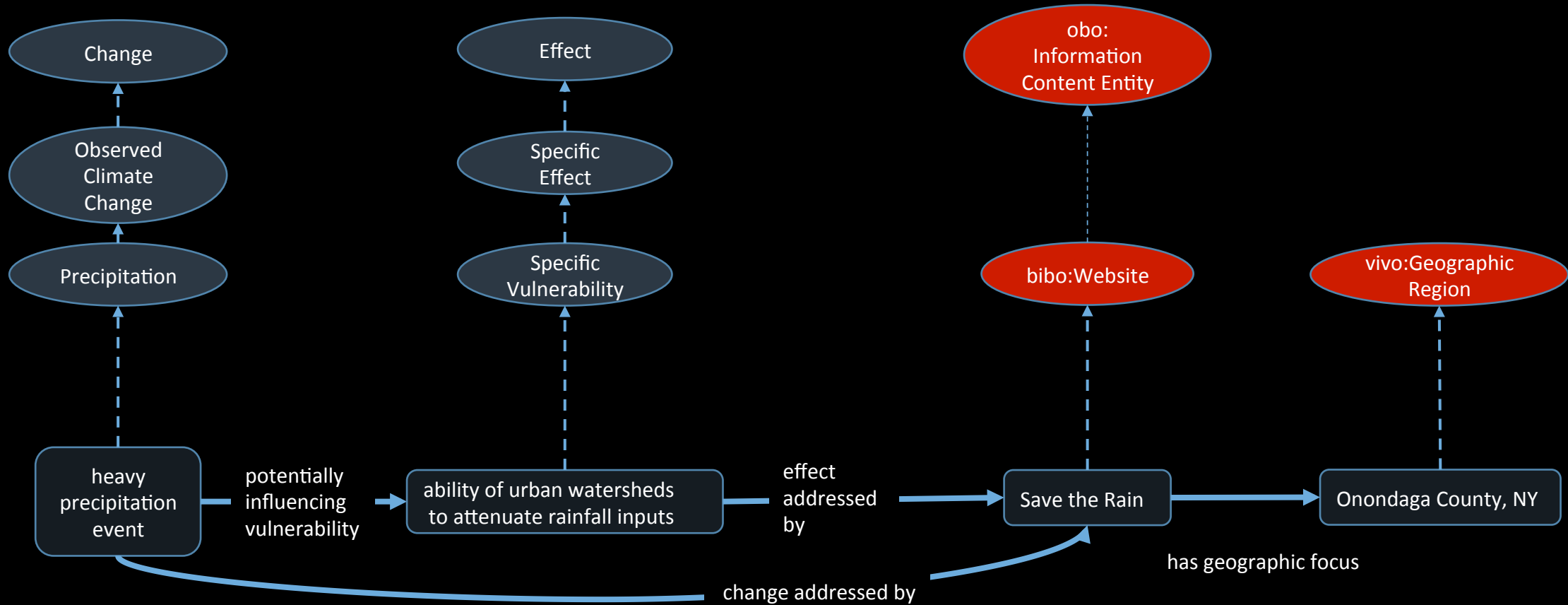
- Leverages existing VIVO-ISF classes (Red)
- Incorporates new classes representing key climate change data elements
- Utilizes existing classgroups and adds new classgroups that become facets on the front end



# Search and Display



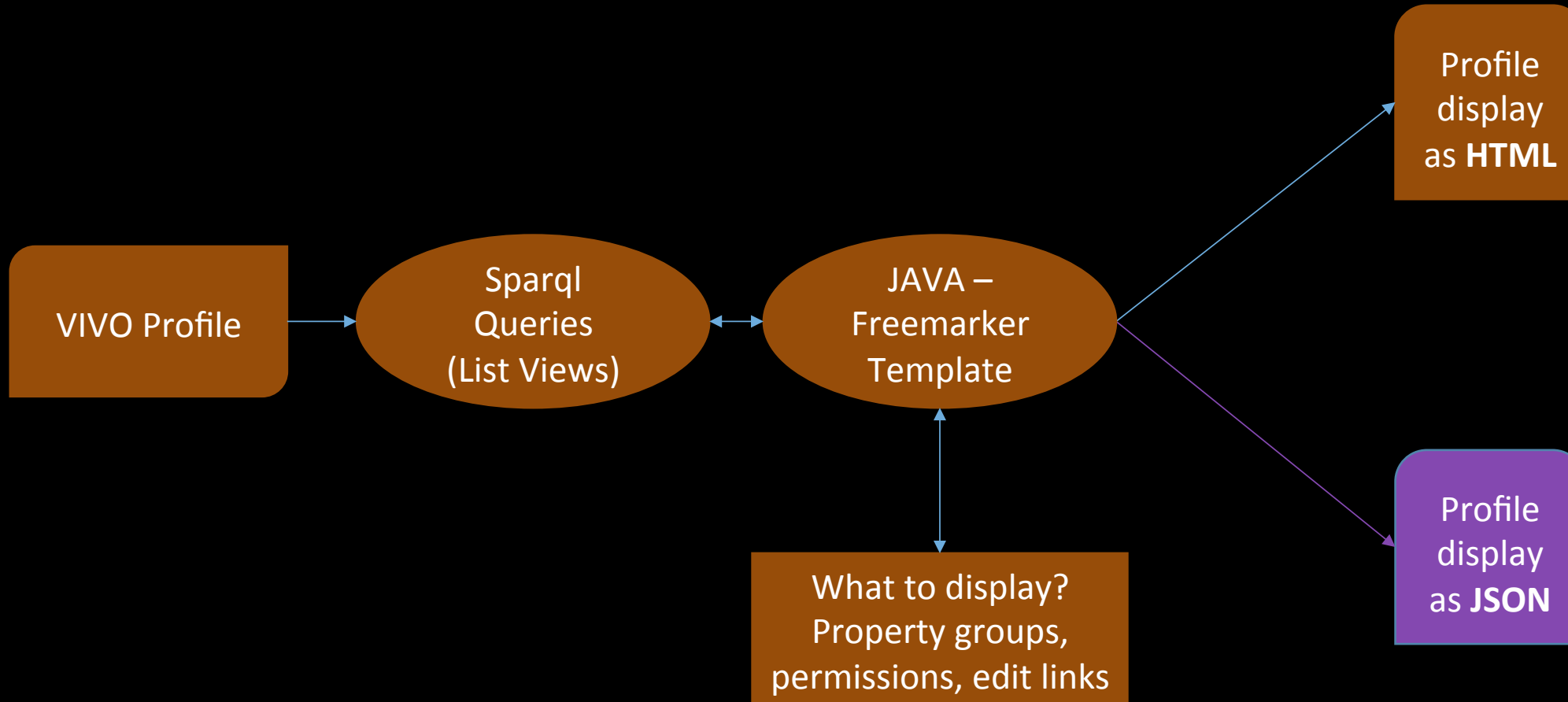
# Translating relationships to search



# PostGIS: Geographical boosting

- Custom solr function
  - Input = Bounding Box coordinates
  - PostGIS API: Retrieve vivo URIs related to bounding box
    - Overlap/closely approximate bounding box
  - Output = rank in PostGIS results
- Blacklight: Solr boosts by custom function
  - Items closer to/overlapping bounding box are boosted in search results

# VIVO Profile as JSON



# Reflections and Future Work

- NYCCSC ontology as linked open data
- Community input from larger VIVO-ISF ontology community
- Expanding the clearinghouse
  - E.g. Single VIVO instance for metadata, multiple front-ends
  - Multiple states
  - Regional clearinghouse
  - Incorporating additional linked data
    - DBPedia, GeoNames

# Questions