# Adaptable Information Models in the Global Change Information System

Brian Duggan[12], Andrew Buddenberg[3],
Steve Aulenbach[12], Robert Wolfe[14], Justin Goldstein[12]

[1]US Global Change Research Program
[2]University Coporation for Atmospheric Research
[3]National Oceanic and Atmospheric Administration
[4]National Aeronautics and Space Administration

December 16, 2014

`http://data.globalchange.gov`
`http://github.com/USGCRP/gcis`

# Outline

1. Introduction and Functionality
    - NCA3 Report Production
    - NCA3 Website backend
    - Provenance of Resources
    - Source of reliable information
    - Connections with other sources of information

2. Information Model
    - Relational
    - Semantic
    - Examples

3. System Architecture
    - Diagram
    - Updating the Schema
    - Updating the Ontology
    - Updating Content

In May, 2014, the US Global Change Research Program released the 2014 National Climate Assessment.

Producing this 829 page report and its web site involved collaboration between over 300 authors, numerous editors, graphics producers, scientists, data scientists, software developers, and web teams.

The content included 161 findings, 284 figures, 3,395 bibliographic references (journal articles, books, reports).

Function 1 : support the production of the third National Climate Assessment.
One year earlier, in May 2013, the USGCRP had the first public release of the Global Change Information System.

The GCIS provides identifiers and a RESTful API for resources used in the construction of the report.
http://data.globalchange.gov

- /report/nca3
- /report/nca3/finding/extreme-precipitation-increase
- /article/10.1080/15287390801997625

# The Third National Climate Assessment

This 800 page document and its website involved **colloboration** between 300 authors, numerous editors, graphics producers, scientists, data scientists, software developers, and web teams.

# collaboration

The GCIS facilitated the assembly of the report by providing common **identifiers** for resources and concepts.

# identifiers

Identifiers are URIs and correspond to explicitly defined concepts. They can be read or written using a **RESTful API**.

# RESTful API

The architecture for the GCIS is built around providing :

- a RESTful API
  GET a URL for JSON, Turtle or HTML
- Triple store
  URIs in the triple store are resolvable URLs in the API.

# architecture

ingest - POST/PUT - relational database
– templates – turtle - **triple store**
– JSON - API/faceted search

# SPARQL

## http://bit.ly/gcis-dbpedia

```
PREFIX bibo: <http://purl.org/ontology/bibo/>
PREFIX gcis: <http://data.globalchange.gov/gcis.owl#>
PREFIX cito: <http://purl.org/spar/cito/>
PREFIX dcterms: <http://purl.org/dc/terms/>
PREFIX dbprop: <http://dbpedia.org/property/>
PREFIX dbpo: <http://dbpedia.org/ontology/>

SELECT  DISTINCT ?dbpjournal ?gcisjournal ?issn
FROM <http://data.globalchange.gov>
WHERE {
    SERVICE <http://data.globalchange.gov/sparql> {
        ?gcisjournal a bibo:Journal .
        ?gcisjournal bibo:issn ?issn .
        ?gcisjournal dcterms:hasPart ?gcisarticle .
        ?gcisarticle a bibo:Article .
        ?gcisarticle dcterms:isPartOf ?gcisjournal .
        ?gcisarticle cito:isCitedBy <http://data.globalchange.gov/report/nca3> .
     }
   SERVICE <http://dbpedia.org/sparql> {
    ?dbpjournal dbprop:frequency "Monthly"@en .
    ?dbpjournal dbpo:issn ?issnd .
    }
  FILTER(?issnd = ?issn)
}
```

# results

go here

# Resources

GCIDs

http://data.globalchange.gov

- /article/10.1080/15287390801997625
- /report/usfs-pnw-gtr-855
- /reference/007a7014-723e-4ceb-a395-5c986b1bf884
- /report/nca3/figure/global-temperature-and-carbon-dioxide
- /image/26fc56f4-b4e0-425b-adc8-14c6d961d558
- /report/nca3/table/decisions-scales
- /report/nca3/finding/extreme-precipitation-increase
- /organization/nasa
- /person/0000-0001-6667-7047
- /dataset/nca3-cddv2-r1

# Functionality

- Support NCA3 report production
- Support NCA3 website (client side jQuery)
- Provide minimal landing pages for resources
- Provide a public JSON API
  `http://data.globalchange.gov/api_reference`
- Provide semantic information
- Be interoperable (e.g. use existing identifiers)
- Provide a public SPARQL endpoint
  `http://data.globalchange.gov/sparql`
- JSON, RDF, schema.org, HTML, Turtle, RDF-XML

# Testing

- Test driven development (unit tests)
- SPARQL tests
- Continuous Integration Testing (github, travis-ci.org)
- Test driven data acquisition
- Continuous Content Validation
  `http://github.com/USGCRP/gcis-qa`

# Server Architecture

- RDBMS (PostgreSQL) for storage
  Fine-grained transaction auditing, referential integrity
- HTML templates
- Turtle templates (and other formats)
- Scrape into triple store (Virtuoso)
- Data structures into JSON, YAML
- nginx reverse proxy cache

# Clients

- Python (Andrew)
  `http://github.com/USGCRP/gcis-py-client`
- Perl
  `http://github.com/USGCRP/gcis-pl-client`
- Javascript (jQuery)
- php (Drupal)

# Narrative vs structure

# Semantic vs Relational

# Resources

# Identifiers

# Publications, Contributors (Entities, Agents, Activities)

http://data.globalchange.gov/resources

# Discussion