

The Global Change Information System

Nuts and Bolts

Brian Duggan¹², Steve Aulenbach¹²,
Robert Wolfe²³, Justin Goldstein¹²

¹UCAR, ²USGCRP, ³NASA

July 9, 2014

<http://data.globalchange.gov>
<http://github.com/USGCRP/gcis>

Outline

1 NCA3 Report Assembly

- What
- Who
- How
- Resources
- Resources
- Role of the GCIS

2 Implementation

- Functionality
- Architecture
- Server
- Clients
- SPARQL
- Versioning
- Testing

3 Information Model

- Concepts
- Details

4 Discussion

What

- The Third National Climate Assessment
- PDF(s)
- Website
- Information System

Who

- Scientists (Authors)
- Science analysts
- Editors
- Graphic designers
- Web developers
- Project Managers

How

- Spreadsheets
- Google docs
- Email
- Endnote
- Scientific Software
- Graphics Software
- Content Management Systems
- Wikis
- Various miscellaneous desktop and cloud software

Resources

The tools are used to represent and manipulate various resources.

- Journal Articles
- Reports
- References
- Figures
- Images
- Tables
- Findings
- Organizations
- People
- Datasets

Resources

<http://data.globalchange.gov>

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

Role of the GCIS

- Common points of reference
- Common vocabulary across teams
- Language, terminology, vocabulary, ontology
- Uniform Resource Identifiers
- URLs are actionable : URLs
- Information manipulation via API or web forms
- Information extraction via API or browsing
- Information modeling with relational or semantic models
- Fine grained tracking of all changes.
- Convenient useful information entry
- Highly scalable information retrieval

Functionality

- Support nca2014.globalchange.gov (jQuery)
- GCIS provides JSON backend
- Figures have client side calls
- Also uses dataset landing pages
- API used for data ingestion and retrieval
- Support semantic queries
- SPARQL endpoint
`http://data.globalchange.gov/sparql`
- JSON, RDF, RDF-A, HTML, Turtle, RDF-XML

Architecture

- RDBMS (PostgreSQL) for storage
- HTML templates
- Turtle templates into other formats
- Scrape into triple store
- Data structures into JSON, YAML

Server

- Perl (mojolicious)
- nginx, proxies
- postgres
- Virtuoso
- Caching

Clients

- Python (Andrew)
- Perl
- Javascript (jQuery)
- php (Drupal)

SPARQL

<http://bit.ly/1ilgeQz>

```
PREFIX dbpediaowl: <http://dbpedia.org/ontology/>
PREFIX bibo: <http://purl.org/ontology/bibo/>
PREFIX gcis: <http://data.globalchange.gov/gcis.owl#>
PREFIX cito: <http://purl.org/spar/cito/>
```

```
SELECT DISTINCT ?gcisjournal
FROM <http://data.globalchange.gov/sparql>
WHERE
```

```
    SERVICE <http://data.globalchange.gov/sparql>
      ?gcisjournal a bibo:Journal .
      ?gcisjournal bibo:issn ?issn .
      ?gcisarticle gcis:inPublication ?gcisjournal .
      ?gcisarticle cito:isCitedBy <http://data.globalchange.gov/report/nca3> .
```

```
    BIND(STRLANG(?issn, "en") AS ?issn_en)
```

```
    SERVICE <http://dbpedia.org/sparql>
      ?dbpjjournal dbpediaowl:frequencyOfPublication "Monthly"@en .
      ?dbpjjournal dbpediaowl:issn ?issn_en .
    FILTER(STR(?issn_en) = ?issn)
```

Versioning

- git for code
- Postgres audit mechanisms for content

Testing

- Test driven development (unit tests)
- SPARQL tests
- Continuous Integration Testing
- Test driven data acquisition
- Continuous Content Validation (gcis-qa)

Narrative vs structure

Semantic vs Relational

Resources

Identifiers

Publications, Contributors

<http://data.globalchange.gov/resources>

Discussion