The Global Change Information System
Nuts and Bolts

Brian Duggan[1,2], Steve Aulenbach[1,2],
Robert Wolfe[2,3], Justin Goldstein[1,2]

[1]UCAR, [2]USGCRP, [3]NASA

July 9, 2014

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

Outline

NCA3 Report Assembly
Implementation
Information Model
Discussion

What
Who
How
Resources
Resources
Role of the GCIS

What

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

NCA3 Report Assembly
Implementation
Information Model
Discussion

What
Who
How
Resources
Resources
Role of the GCIS

Who

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

NCA3 Report Assembly
Implementation
Information Model
Discussion

What
Who
How
Resources
Resources
Role of the GCIS

How

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

NCA3 Report Assembly
Implementation
Information Model
Discussion

What
Who
How
Resources
Resources
Role of the GCIS

Resources

The tools are used to represent and manipulate various resources.

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

NCA3 Report Assembly
Implementation
Information Model
Discussion

What
Who
How
Resources
Resources
Role of the GCIS

Resources

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

NCA3 Report Assembly
Implementation
Information Model
Discussion

What
Who
How
Resources
Resources
Role of the GCIS

Role of the GCIS

- Common points of reference
- Common vocabulary across teams
- Language, terminology, vocabulary, ontology
- Uniform Resource Identifiers
- URIs 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

NCA3 Report Assembly
Implementation
Information Model
Discussion

Functionality
Architecture
Server
Clients
SPARQL
Versioning
Testing

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

NCA3 Report Assembly
Implementation
Information Model
Discussion

Functionality
Architecture
Server
Clients
SPARQL
Versioning
Testing

Architecture

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

NCA3 Report Assembly
Implementation
Information Model
Discussion

Functionality
Architecture
Server
Clients
SPARQL
Versioning
Testing

Server

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

NCA3 Report Assembly
Implementation
Information Model
Discussion

Functionality
Architecture
Server
Clients
SPARQL
Versioning
Testing

Clients

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

NCA3 Report Assembly
Implementation
Information Model
Discussion

Functionality
Architecture
Server
Clients
SPARQL
Versioning
Testing

SPARQL

`http://bit.ly/gcistest`

```
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>
     ?dbpjournal dbpediaowl:frequencyOfPublication "Monthly"@en .
     ?dbpjournal dbpediaowl:issn ?issn_en .
     FILTER(STR(?issn_en) = ?issn)
```

NCA3 Report Assembly
Implementation
Information Model
Discussion

Functionality
Architecture
Server
Clients
SPARQL
Versioning
Testing

Versioning

- git for code
- Postgres audit mechanisms for content

NCA3 Report Assembly
Implementation
Information Model
Discussion

Functionality
Architecture
Server
Clients
SPARQL
Versioning
Testing

Testing

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

NCA3 Report Assembly
Implementation
Information Model
Discussion

Concepts
Details

Narrative vs structure

NCA3 Report Assembly
Implementation
Information Model
Discussion

Concepts
Details

Semantic vs Relational

NCA3 Report Assembly
Implementation
Information Model
Discussion

Concepts
Details

Resources

NCA3 Report Assembly
Implementation
Information Model
Discussion

Concepts
Details

Identifiers

NCA3 Report Assembly
Implementation
Information Model
Discussion

Concepts
Details

Publications, Contributors

NCA3 Report Assembly
Implementation
Information Model
Discussion

Concepts
Details

http://data.globalchange.gov/resources

Discussion