NOAA
Steps to the Geoverse
Shaping the Next Generation of Earth Systems Compute

Ryan Berkheimer, NOAA
Agenda ID: 2023.04.18_15.10

WGISS-55
Cordoba, Argentina (CONAE)
18-20 April 2023
Executive Summary

- The UN-GGIM in the July 2022 Discussion Paper on the Geoverse recognizes the significance of a distributed and machine dominated world
  - It is understood that disruptive circumstances require disruptive solutions
- NOAA is currently developing a holistic capability to enable the Geoverse System of Systems (SoS) step change
  - Cloud based agile reference architecture for evolutionary governance and open innovation
  - Iterative inclusion and improvement of existing systems
- The capability is being driven by many requirements and considerations
  - Web 3.0 standards vision for full semantic interoperability - achieving the original vision of the WWW
  - Full definition and process provenance for trust and scientific reproducibility - treating process as data
  - Support of federated, multi-owner earth systems digital twins and sensor networks
- The capability is being pursued through many intra-agency and inter-agency partnerships
  - Interagency Study Groups - NOAA/NASA/CNES Coastal Zone Digital Twins
  - NSF - Open Knowledge Network (‘next-gen public data infrastructure similar in scale to the internet’); the Public/Private I/UCRC
  - Strategic Alignment - NOAA’s Societal Data Insights Initiative, PARR 2.0
  - Industry and Standards - OGC Disaster Pilot; Climate Resiliency Pilot; Testbed 19 for Interoperability and Agile Architecture
A Big Idea - Let’s Build a Runnable, Trustable, Democratic Internet to Improve the Earth System

Figure 4. From data to information, knowledge and wisdom. Adapted from DIKW Model for knowledge management and data value extraction.
Visualizing a Step Change to the Geoverse

- A democratized system of systems powered by machine to machine communication
  - Machines as users
  - Humans as users
- An open ecosystem in which all users can both consume and contribute information
- A federated framework of universally useful understanding

Figure 3. The future geospatial information ecosystem comprising SDIs, SoS and the Geoverse. [Future Geospatial Information Ecosystem: From SDI to SoS and on to the Geoverse]
Why Haven’t Spatial Data Infrastructures Been Enough?

- **At every level** of the global spatial data infrastructure, our system is characterized by data silos
  - Data access is generally difficult
  - Usage is generally difficult
  - Collaboration is generally difficult
  - Operationalizing is generally difficult
  - Universal Standards management is nearly impossible - so how can we merge things?
  - The Big Idea is nearly impossible to achieve under this condition

- **Data Silos are expensive and limiting**
  - Resource overhead in duplicating efforts in feature sets, maintenance, management, etc.
  - Enormous costs incurred due to resolving conflicts of definition, provenance, use rights, etc.
  - Service offerings that don’t match data-driven user preferences
  - Inability to do holistic comparison of competitive approaches and/or products
  - Lack of holistic quality controls
  - Reproducibility crisis leading to trust issues

- **This problem isn’t limited to any particular enterprise**
  - Data integration in earth science is a global issue
  - Enormous value in enabling connection between organizations - global system of systems
  - No one organization can solve this problem alone!!! It requires an orchestrated federation of players

User Perspective

A critical element of successful disaster management is collaboration between stakeholders such as represented by these personas, both through sharing of data / information, decision on useful indicator recipes, and through direct exchange of knowledge that leads to better ideas and actions.

OGC Disaster Pilot 2023

We aren’t here quite yet
Using the Federated Global Infrastructure Framework (FGIF) as a guide and ‘unit test suite’

- Geoverse enablement is a shared **social** problem
  - SDI methods have moved us forward
  - Geoverse is immensely more complex
  - A federated problem across boundaries

- The FGIF provides a high level framework for alignment of verification and validation of solutions by enabling semantic interoperability in communications
  - Highest level of understanding
  - Rises above agency and national solutions
  - Shared understanding of goals to shape
    - Policy
    - Solutions

- Treat the FGIF as a leveling framework
  - For a given solution we can ask -
    - how mature is it?
    - In what ways should it evolve to become more mature?

**Future Geospatial Information Ecosystem: From SDI to SoS and on to the Geoverse**
A full resolution earth is a system of systems
Lots of humanity spanning effort to measure, understand, and explain
Data is fundamentally interoperable due to scientific frameworks of understanding
But - most holistic interoperability efforts fail in some way - why?

Conway’s Law: any organization that designs a system (defined broadly) will produce a design whose structure is a copy of the organization’s communication structure
  ○ Conway’s Law implies hard limits to large efforts in terms of syntactic, schematic, semantic, and legal interoperability constraints

The FGIF gives us a shared vision and validation framework at the global level
- Needs lower level frameworks to address verification and validation of implementation

To this end, NOAA created the Digital Twin Earth Framework Specification (DTE-FS)
- DTE-FS is a first-principles derivation of specific actionable requirements for validating a framework supportive of an interoperable, federated, & holistic earth systems digital twin

Requirements for Federated Digital Twins as described by the DTE-FS represent an operational blueprint to practical implementation of the SoS step change toward the Geoverse
- Categorized and leveled interoperability
  - Accommodation for Conway’s Law
- Functional requirements to address uses and goals

NOAA has been implementing the DTE-FS as the Next Generation Archive and Access System
- DTE-FS mapped and tailored to specific requirements
  - National Open Knowledge Network and the Semantic Web (5 star linked open data); NARA (full provenance) and PARR 2.0 (open science); FAIR and CARE data
The DTE-FS in One Image

Open Architecture Specification

- Architecture Driven by **Interoperability**
- Open Architecture - All architectural components must be changeable and flexible
- Architecture components must be segmentable
  - Scalability demands it!!!

Reference Model and Requirements
I can create processes
I can share processes
I can discover processes
I can inspect processes
I can evolve processes
I can copy processes
I can link processes
I can process shop
My machine can do all these things, too!

Critical Understanding #1: **Process as a Dataset**

- Process is ‘space complete’
- We can describe any system as a set of processes
- Process-as-data enables syntactic interoperability
  - Request/Response/Records

DTE User System
Illustrative Example - Dynamic Process Provenance
Critical Understanding #2: Semantic Interoperability

- Process-as-data alone cannot solve interoperability
- Semantic interoperability within a shared reference model is required to enable automated terminology resolution, value transformation, and evolutionary governance
- Critically, Semantic interoperability enables higher order legal interoperability automation
- Critically, Semantic interoperability enables decision definition and capture, enabling automated feedback and introduces observability to system improvement

Archetypes: Constraint-based Domain Models for Future-proof Information Systems

ISO 13606

The basic building block - recursive and generic

A complete generic model for DTE-FS data
Illustrative Example - Fuzzy Semantics
Critical Understanding #3: **Process as a Digital Thread**

- Structured KOS
- FAIR
  - Findable
  - Accessible
  - Interoperable
  - Reusable
- CARE
  - Fine-grain access rights
- Knowledge Graph held
- Independent spaces
- API Discovery

- Process as a digital thread allows us to compose identity, transformation, and delivery tasks that can use one or more of any preexisting, well-defined ontologies available on the semantic web (e.g. SSN, SOSO, Schema.org, SensorML, Dublin Core, GCMD, etc.), and/or make use of newly minted, user defined ontologies
- Process as a digital thread enables root cause analysis
- Process as a digital thread enables full provenance for reproducibility
- Process as a digital thread enables dynamic ‘at a distance’ composability
- Process as a digital thread enables process shopping
  - Pre-flight cost estimation via describe plans
- Process as a digital thread enables evolutionary plug-and-play
Illustrative Example - Process as a Digital Thread
Critical Understanding #4: Legal Interoperability

- Legal interoperability deals with ability to
  - discover other processes
  - understand access rights of processes
  - understand payment requirements for using processes
  - use processes and pay for them automatically
- Approaching legal interoperability in this way, in combination with a process-as-data view, allows us to federate globally
- Legal interoperability allows us to democratize existing funding models
  - More agile
  - Better results
  - Credit where credit is due
  - Fair use
  - Trusted use

![Access Rights Policy Diagram](image)
Tangible Takeaways - What Are We Enabling?

We are creating an architecture to host the emerging generation of disruptive technology in a way that harnesses it for trustworthy science and makes it available to all.

We are enabling new insights to answer our most challenging problems.

ChatGPT Gets Its “Wolfram Superpowers”!

March 23, 2023

ChatGPT Gets Its “Wolfram Superpowers”!

March 23, 2023

ChatGPT Gets Its “Wolfram Superpowers”!

March 23, 2023
What’s next?

- NOAA continues to socialize and coordinate integration with the framework in a controlled environment (the NESDIS Common Cloud Framework)
  - Refining implementation details
  - Assessing and improving performance
- Partnership contexts are refining and validating higher order interoperability and functional requirements
- Continuously mapping back to Geoverse goals and milestones to get a sense of gaps and timelines
- Tackling problem of legal interoperability using widespread collaborations
Reference and Backups
The following slides contain the pulled-out tables from the UN-GGIM Future Global Infrastructure (FGIF) Discussion Paper. Tables contain step changes and gap analysis tools for use in determining if the geoverse is being enabled or how to approach enablement in the case of identified gaps.
### Organizational Migrations - Governance/Institutions

<table>
<thead>
<tr>
<th>Strategic Pathway Elements</th>
<th>Spatial Data Infrastructures</th>
<th>Future Geospatial Information Ecosystem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leadership</td>
<td>Strategies and plans on geospatial data management</td>
<td>Global interconnectivity to leverage knowledge creation opportunities</td>
</tr>
<tr>
<td>Institutional Arrangements</td>
<td>Traditional hierarchical SDI governance arrangements</td>
<td>Agile non-hierarchical inclusive multi-actor governance configurations</td>
</tr>
<tr>
<td>Value Proposition</td>
<td>Access to fundamental data</td>
<td>Affordable and equitable access to knowledge</td>
</tr>
<tr>
<td>Governance Model</td>
<td>SDI data governance models</td>
<td>Global Geospatial Knowledge Governance Framework</td>
</tr>
</tbody>
</table>

Figure 7. Governance and Institutions Strategic Pathway – the Step Change

Source: UN-GGIM July 2022 Discussion Paper on Future Data Systems
## Organizational Migrations - Policy/Legal

### Strategic Pathway Elements
- **Legislation**
  - National policy and legal frameworks for the democratization of data
  - Global guidance for national policy and legal frameworks for the democratization of knowledge

### Spatial Data Infrastructures
- **Policies, Norms and Guides**
  - Standalone geospatial policies and laws for data management, access and use.
  - Geospatial policies and laws interoperable with wider government digital policy and global knowledge management agendas

### Future Geospatial Information Ecosystem
- **Governance and Accountability**
  - A reliance on humans to comply with geospatial data policies and laws
  - Use of technology to enforce geospatial knowledge management and use, as well as internet compliance

- **Data Protection, Licensing and Sharing**
  - Open data policy balanced with protection of sensitive material and respected rights
  - Knowledge creation and the ethical challenges of information bias, digital identities, usage and cybersafety

---

**Figure 8. Policy and Legal Strategic Pathway – the Step Change**

Source: UN-GGIM July 2022 Discussion Paper on Future Data Systems
### Strategic Pathway Elements

<table>
<thead>
<tr>
<th>Strategic Pathway Elements</th>
<th>Spatial Data Infrastructures</th>
<th>Future Geospatial Information Ecosystem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business Model</td>
<td>Traditional bi-directional Purchaser/Provider Models</td>
<td>Decentralized autonomous business models enabled by advanced internet technologies</td>
</tr>
<tr>
<td>Opportunities</td>
<td>Creation of geospatial data to support the delivery of new products and services</td>
<td>Creation of geospatial analytics to support knowledge on-demand</td>
</tr>
<tr>
<td>Investment</td>
<td>Investment in geospatial data and interfaces for SDI programs and projects</td>
<td>Investment in machine-readable data, AI, blockchain etc. enabling participation in global digital connectivity</td>
</tr>
<tr>
<td>Benefits Realization</td>
<td>Economic development and societal benefits derived from geospatial products and services</td>
<td>New revenue and cost savings from tokenized multimodal business opportunities</td>
</tr>
</tbody>
</table>

*Figure 9. Financial Strategic Pathway – the Step Change*

*Source: UN-GGIM July 2022 Discussion Paper on Future Data Systems*
Organizational Migrations - Data

Figure 10. Data Strategic Pathway – the Step Change

Source: UN-GGIM July 2022 Discussion Paper on Future Data Systems
## Organizational Migrations - Innovation

### Strategic Pathway Elements

<table>
<thead>
<tr>
<th>Technological Advances</th>
<th>Spatial Data Infrastructures</th>
<th>Future Geospatial Information Ecosystem</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Digitization and diversification of data sources, and cloud computing</td>
<td>4IR technologies, Internet, geoanalytics, semantic tools, sensor networks, AI and edge computing</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Innovation and Creativity</th>
<th>Business innovation: data, products and services</th>
<th>Social innovation: knowledge on-demand and ethical human-centered design.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Process Improvement</th>
<th>Digital transformation, increased productivity and improved data quality</th>
<th>Enriched knowledge economies, responsive communications and enhanced reliability of knowledge</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Bridging the Digital Divide</th>
<th>G2G relationships that deliver hardware, software and knowhow</th>
<th>Knowledge sharing that delivers inclusivity and digital equity for communities</th>
</tr>
</thead>
</table>

**Figure 11. Innovation Strategic Pathway – the Step Change**

**Source:** UN-GGIM July 2022 Discussion Paper on Future Data Systems
<table>
<thead>
<tr>
<th>Strategic Pathway Elements</th>
<th>Spatial Data Infrastructures</th>
<th>Future Geospatial Information Ecosystem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standards Governance and Policy</td>
<td>A commitment to assess, establish, and maintain a common standards framework</td>
<td>A commitment to adopt standard models for knowledge representation</td>
</tr>
<tr>
<td>Technology and Data Interoperability</td>
<td>Voluntary standards for data and technology</td>
<td>Anticipatory regulations, FAIR data and human-centered standards for knowledge creation</td>
</tr>
<tr>
<td>Compliance Testing and Certification</td>
<td>Regular assessments, training, government mandates, performance metrics, testing and certification</td>
<td>Regulatory sandboxes, proof of concepts and incentives for deploying and scaling 4IR technologies</td>
</tr>
<tr>
<td>Community of Practice</td>
<td>Sharing and leveraging data and technology standards</td>
<td>Inclusive and participatory roles in data vocabularies, digital twins, and other 4IR standards development</td>
</tr>
</tbody>
</table>

Figure 12. Standards Strategic Pathway – the Step Change

Source: UN-GGIM July 2022 Discussion Paper on Future Data Systems
<table>
<thead>
<tr>
<th>Strategic Pathway Elements</th>
<th>Spatial Data Infrastructures</th>
<th>Future Geospatial Information Ecosystem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross Sector and Interdisciplinary Cooperation</td>
<td>Investment, data and resource sharing</td>
<td>New horizons for innovation and knowledge creation</td>
</tr>
<tr>
<td>Private Sector and Academia Collaboration</td>
<td>Government led regulated public-private-partnerships</td>
<td>Private and academia sector led decentralized collaborations</td>
</tr>
<tr>
<td>Community Participation</td>
<td>Use of VGI and crowdsourcing to collect local information</td>
<td>Citizen sensing technologies that lead to new innovations and problem solving</td>
</tr>
<tr>
<td>International Collaboration</td>
<td>Strategic partnerships that support national initiatives and agendas</td>
<td>Collaborations that support both national and global innovations</td>
</tr>
</tbody>
</table>

Figure 13. Partnerships Strategic Pathway – the Step Change

Source: UN-GGIM July 2022 Discussion Paper on Future Data Systems
<table>
<thead>
<tr>
<th>Strategic Pathway Elements</th>
<th>Spatial Data Infrastructures</th>
<th>Future Geospatial Information Ecosystem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Awareness</td>
<td>Advocacy and promotion on the value of geospatial information</td>
<td>Education and outreach on the potential of disruptive innovations using integrated data</td>
</tr>
<tr>
<td>Formal Education</td>
<td>Deep subject expertise in surveying, geospatial, geography and geodesy</td>
<td>Foundational concepts in computer and Internet science and 4IR information technologies</td>
</tr>
<tr>
<td>Professional Training</td>
<td>Lifelong learning and development through hands-on experiences</td>
<td>Intensive development in new and emerging paradigms</td>
</tr>
<tr>
<td>Entrepreneurship</td>
<td>Capacity development leading to innovation and new business ventures</td>
<td>Collaborative innovation hubs that stimulate growth in interconnected products, value chains and business models</td>
</tr>
</tbody>
</table>

Figure 14. Capacity and Education Strategic Pathway – the Step Change

Source: UN-GGIM July 2022 Discussion Paper on Future Data Systems
<table>
<thead>
<tr>
<th>Strategic Pathway Elements</th>
<th>Spatial Data Infrastructures</th>
<th>Future Geospatial Information Ecosystem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stakeholder and User Engagement</td>
<td>Narrow focus on government user engagement</td>
<td>Broad spectrum, diverse and inclusive</td>
</tr>
<tr>
<td>Strategic Messaging</td>
<td>Limited brand awareness, complicated messaging and inconsistent terminology</td>
<td>Unified brand and terminology, and clear messaging that sparks conversations and imagination</td>
</tr>
<tr>
<td>Communication Strategy, Plans and Methods</td>
<td>Project and program focussed communications</td>
<td>Strategic, targeted and impactful communications that keep pace with changing times</td>
</tr>
<tr>
<td>Monitoring and Evaluation</td>
<td>Surveys and metrics on effectiveness and efficiency of communications</td>
<td>Knowledge analytics and metrics on participation rates, understanding and change in usage</td>
</tr>
</tbody>
</table>

Figure 15. Communication and Engagement Strategic Pathway – the Step Change

Source: UN-GGIM July 2022 Discussion Paper on Future Data Systems
NOAA System Details
Web Crawler - JSON-LD

Web Crawler Endpoint
https://ncei.nesdis.noaa.gov/archive/{{namespace}}/{{type}}/{{thing_id}}/

Example:
https://ncei.nesdis.noaa.gov/archive/records/aiu/abc_123.jsonld (full document)

context: {"http://rdf.org",...}
body: { "pattern": "http://ncei.nesdis.noaa.gov/archive/patterns/aiu/xyz_123.jsonld", "template": "...", "type": "AIU", "description": "...", "content": "...

S3 Metadata Bucket

Namespace = Core
Reference Model Classes
Small and immutable-per-reference-model surface. Recursive classes provide ability to compose inherently interoperable structures

Namespace = Patterns
Unvalued Schema Individuals
Potential for ontologically rich augmentation via overloaded labeling and tagging

Namespace = Templates
Partially Valued Schema Individuals
Serves as a fast-query inference layer at data-stream resolution of descriptions and representation networks

Namespace = Records
Fully Valued Schema Individuals
Mesh layer, high write velocity, fully featured inferencing, rich membership content