Enabling an Iterative, Open Science Transformation to the Geoverse at NOAA with a Federated Knowledge Mesh

CI Compass (https://ci-compass.org/)
Online Webinar
23 April 2024

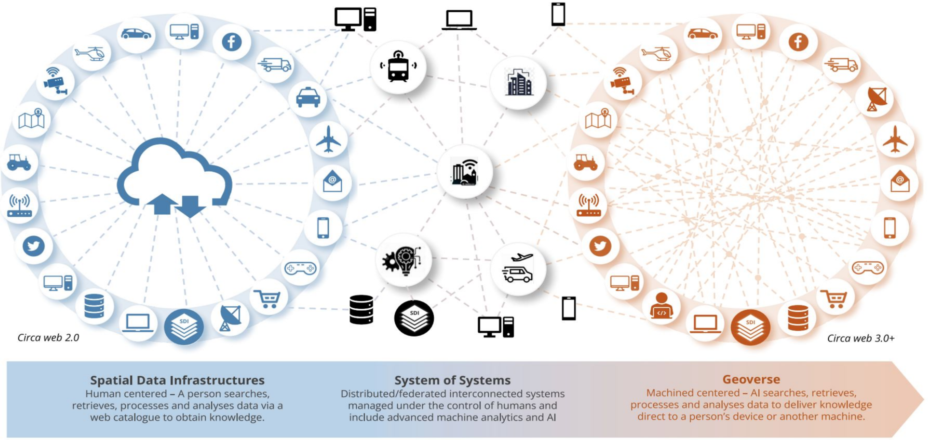
Ryan Berkheimer
Physical Scientist (Position)
Archive Architect (Current Role)
NOAA NESDIS NCEI

National Environmental Satellite, Data, and Information Service

National Centers for Environmental Information

Where To? Geoverse Semantics in Context

- A democratized 'system of systems' powered in large part by machine to machine communication
  - Systems may be 'smart agents'
- An ecosystem in which all users can both consume and contribute information
- A federated trust-based framework for realizing universally useful understanding



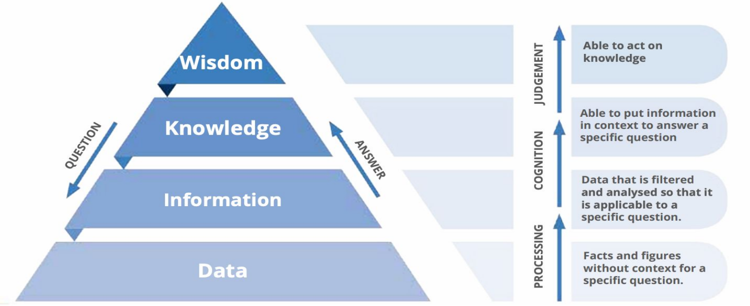


Figure 4. From data to information, knowledge and wisdom. Adapted from DIKW Model for knowledge management and data value extraction.

Strategic Pathway Elements	Spatial Data Infrastructures	Future Geospatial Information Ecosystem
Standards Governance and Policy	A commitment to assess, establish, and maintain a common standards framework	A commitment to adopt standard models for knowledge representation
Technology and Data Interoperability	Voluntary standards for data and technology	Anticipatory regulations, FAIR data and human-centered standards for knowledge creation
Compliance Testing and Certification	Regular assessments, training, government mandates, performance metrics, testing and certification	Regulatory sandboxes, proof of concepts and incentives for deploying and scaling 4IR technologies
Community of Practice	Sharing and leveraging data and technology standards	Inclusive and participatory roles in data vocabularies, digital twins, and other 4IR standards development

Figure 12. Standards Strategic Pathway – the Step Change

What is “Open Science”?

Data access ($$)
Accessible Publications ($$)
Siloed systems
Limited communication
Proprietary Software
“Closed-Tent” culture

Free unlimited data access
Fully documented open software and algorithms
Fully linked data and publications
Open Access Journal publications
Fully Transparent processes
Reproducible across platforms
“Teaching” culture
Open science meetings

FULLY CLOSED — NASA — **FULLY OPEN**

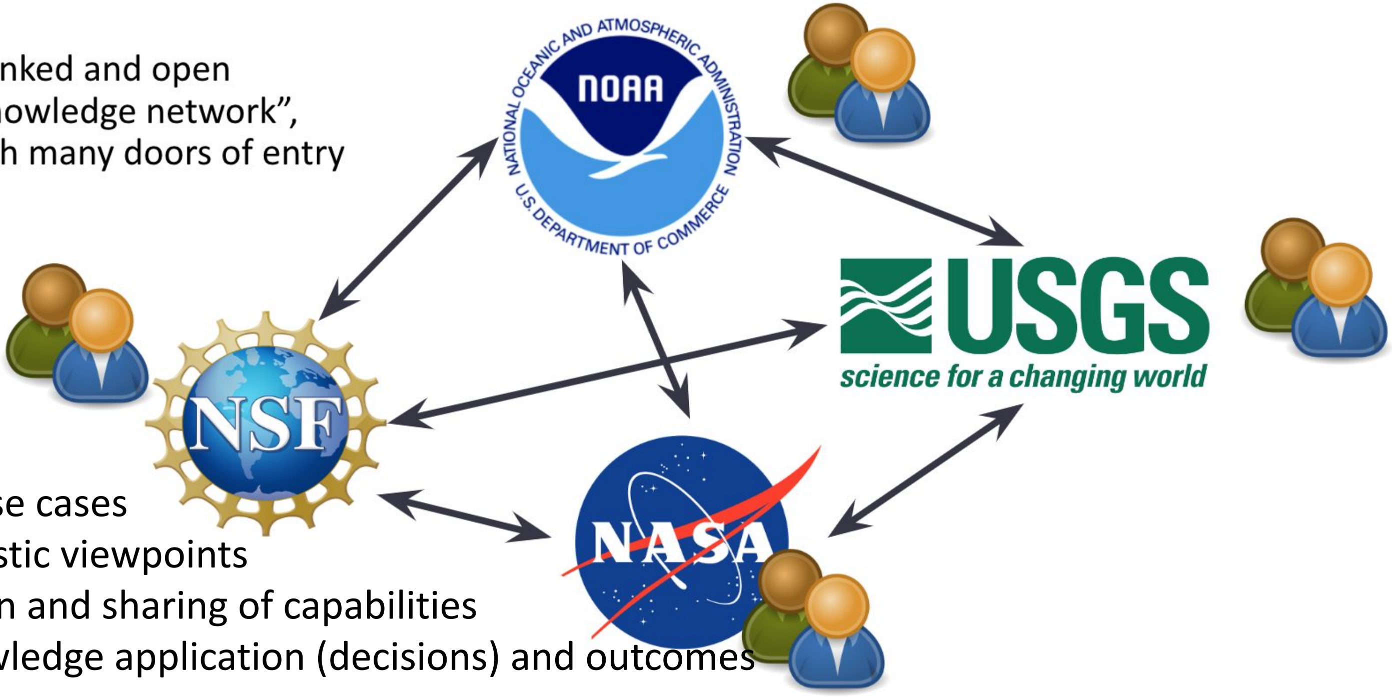
No public access data
No publications
No insight into processes
No reproducibility
“Black Box” culture

Free data access
Open software and algorithms
“Green” Journal publication
Documented processes
Reproducible in specific environments
“Open-Tent” culture

Source: C. Gentemann (2022, NOAA EDM Workshop presentation)

Toward an Open Ecosystem of Trusted Open Ecosystems

A linked and open "knowledge network", with many doors of entry



More coverage in use cases
Providing more holistic viewpoints
Facility specialization and sharing of capabilities
Ability to track knowledge application (decisions) and outcomes

What success means to constituents

Meg, a Coastal Resident and Environmental Enthusiast:

Meg is a resident living on Hilton Head Island with a deep interest in the environment. She talks to the system through her Alexa to gain insights into local marine life, weather patterns, and pollution levels using real-time weather updates, satellite imagery from NASA, and environmental data from NOAA. She can report environmental concerns directly to the EPA with full provenance for investigation.

Bart, a Commercial Fisherman:

Bart relies on accurate and timely information for successful operations in the Gulf. He uses an integrated ship dashboard to make predictions and get updates that help him make informed decisions about optimal fishing locations, avoiding adverse weather conditions, and ensuring sustainable practices. The system enables Bart by combining NOAA's fisheries data, NASA's satellite-based ocean temperature maps, and USGS coastal mapping. Bart sometimes asks the system to explain predictions and give alternative assessments. When needed, Bart further uses the system for direct communication with relevant federal agencies for licensing, regulatory updates, and emergency assistance.

Gertie, an Emergency Response Coordinator:

Gertie requires immediate access to comprehensive information while monitoring wildfires in Colorado. A multi-modal system interface connects them to real-time weather updates, NASA's disaster monitoring, and USGS geological data. Gertie is enabled to coordinate evacuation plans with local and regional authorities and monitor emerging outcomes in real-time.

Holden, a Tourism and Recreation Planner:

Holden aims to create enjoyable experiences for visitors to the Chesapeake Bay. As Holden crafts tailored experiences for his clients, he converses with the system to understand seasonal trends, identify ideal locations, and, shape events accordingly.

Elias, a Public Health Researcher:

Elias is a public health researcher studying coastal communities and relies on diverse data sources. He uses a portal to access and fuse information from health, environmental quality, and satellite record databases to help identify potential health risks and patterns based on air and water quality. He uses context to understand ideal methods of engagement with local communities to provide evidence-based recommendations for improved public health outcomes.

What success means to the workforce

Buffy, a NOAA Archivist:

Buffy needs to efficiently manage and organize a vast number of heterogeneous datasets in the Archive. Through her government portal Buffy accesses and catalogs data. She uses the marketplace to add quality matrix checking, link related datasets, or create new collections and views on the data. She develops, shares, and evolves standard templates and expressions for providers to use and assesses adoption.

Munir, a NOAA Data Manager:

Munir’s job is to coordinate marine data flows between providers and consumers. He uses the system API to help build, monitor, and communicate data pipelines, while ensuring data quality. He uses the API in his own software to achieve on-demand data retrieval in specific formats for generating comprehensive data reports. He contributes his techniques back to the marketplace as reusable templates for use by other data managers.

Marilyn, a NOAA Data Scientist:

Marilyn uses the system daily to access diverse datasets for research and analysis to relate societal indicators to climate change. The system provides here facilitates seamless integration of datasets from NOAA, NASA, USGS, EPA, and NIH. She leverages user and system access data to quickly develop and test new models for improving user experience; the system automatically shares what she creates for reuse by others.

Aditi, a NOAA Physical Scientist:

Aditi builds authoritative climate data records at NOAA. She leverages the system to explore what has been done before, what findings have been made, and what things might need further research. She leverages on-demand data retrieval for existing products in desired formats to try new modeling and simulation techniques. The data and products she produces are made available to the system.

Not pictured: System Administrators; Data Operators; Policy Analysts; etc.

What Does this Mean to the Enterprise?

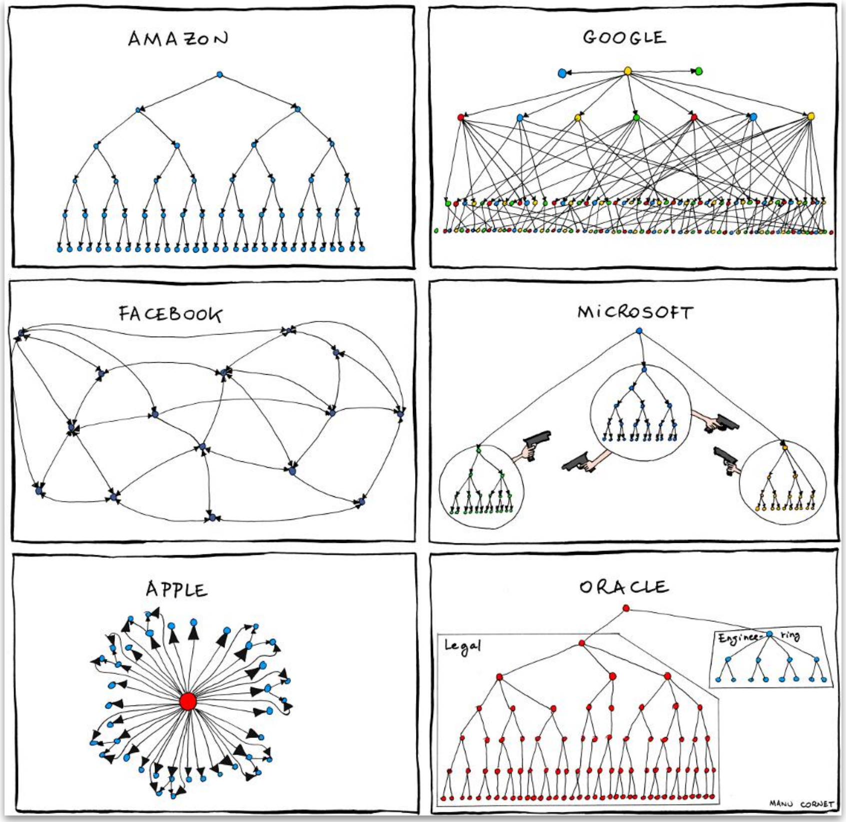
- Working toward a new sustained model
- ["Simple Made Easy" - Rich Hickey (2011)]
  - The model is **simple (decomplected) - gives us maximum concurrency, sustainability**
  - Delivery of the model attempts to make it **easy - tools that are 'near to' users**
- Full provenance of the computational enterprise
  - Data flows, agents, etc
  - Self documenting(!) processes
    - Tracking evolution
    - We can ask the machine
- Shared reference model
  - No more independent stovepipes
  - Iterative improvement (leveling framework)
  - self-sustaining model of shared resources
    - optimizing to use cases, sharing general libraries, etc
- Using advanced tools in the next-gen toolbox
- Enabling new users; user driven discovery; federated interoperability

Supporting the Transformation

”The **architecture** of a **digital system** defines **flexibility**, **capability**, and **functionality**.” - Chaowei Phil Yang, Director and Founder of the NSF Spatiotemporal Innovation Center

Why is Interoperability Needed? A Communications Problem

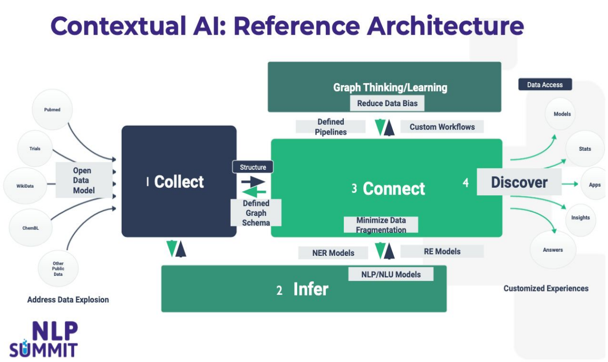
- A full resolution earth is a **system of systems**
- Lots of humanity spanning effort to **measure, understand, and explain**
- **Data is fundamentally interoperable** within scientific frameworks of understanding
- **But -** most holistic interoperability efforts fail in some way - why?
- **Conway's 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
  - **Encoded in SWEBOK 4.0**



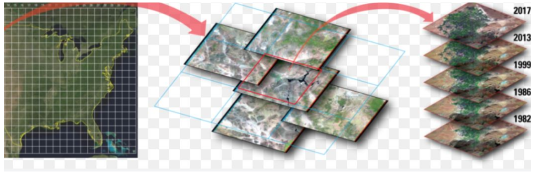
http://scrumbook.org/product-organization-pattern-language/conway-s-law.html

Major Considerations for Digital Twin Federation (Interop) Architecture

- **Agree** on the fact that **Knowledge Sharing is a federated problem**
  - Organizational specializations
  - Domain characteristics
  - Continuous innovations and reconfigurations
  - Competing priorities
  - Nearly infinite user concerns and use cases
- **Consider intra-twin interoperability** concerns, e.g.
  - How do we ensure **reproducibility**? (not science without it!)
  - How how do we **automatically combine** different timescales, units, etc.?
  - How do we **capture provenance** in a FAIR way?
  - How do we **associate uncertainty** with prediction and data?
  - How do we guarantee **performance and scale**?
- **Consider inter-twin interoperability** concerns, e.g.
  - How do we **capture decisions** (non-spatial-data) and associate with lineage and provenance?
  - How do we **avoid limiting** ourselves to a particular **catalog** viewpoint?
  - How do we support **pre-flight cost** simulation, aka **process shopping,** for a given describe plan?
  - How do we **enable inference** that helps us look for unknown connections?
  - How do we **manage legal interoperability** issues, **CARE** issues?
    - *Access rights to processes?*
    - *Bad actors?*

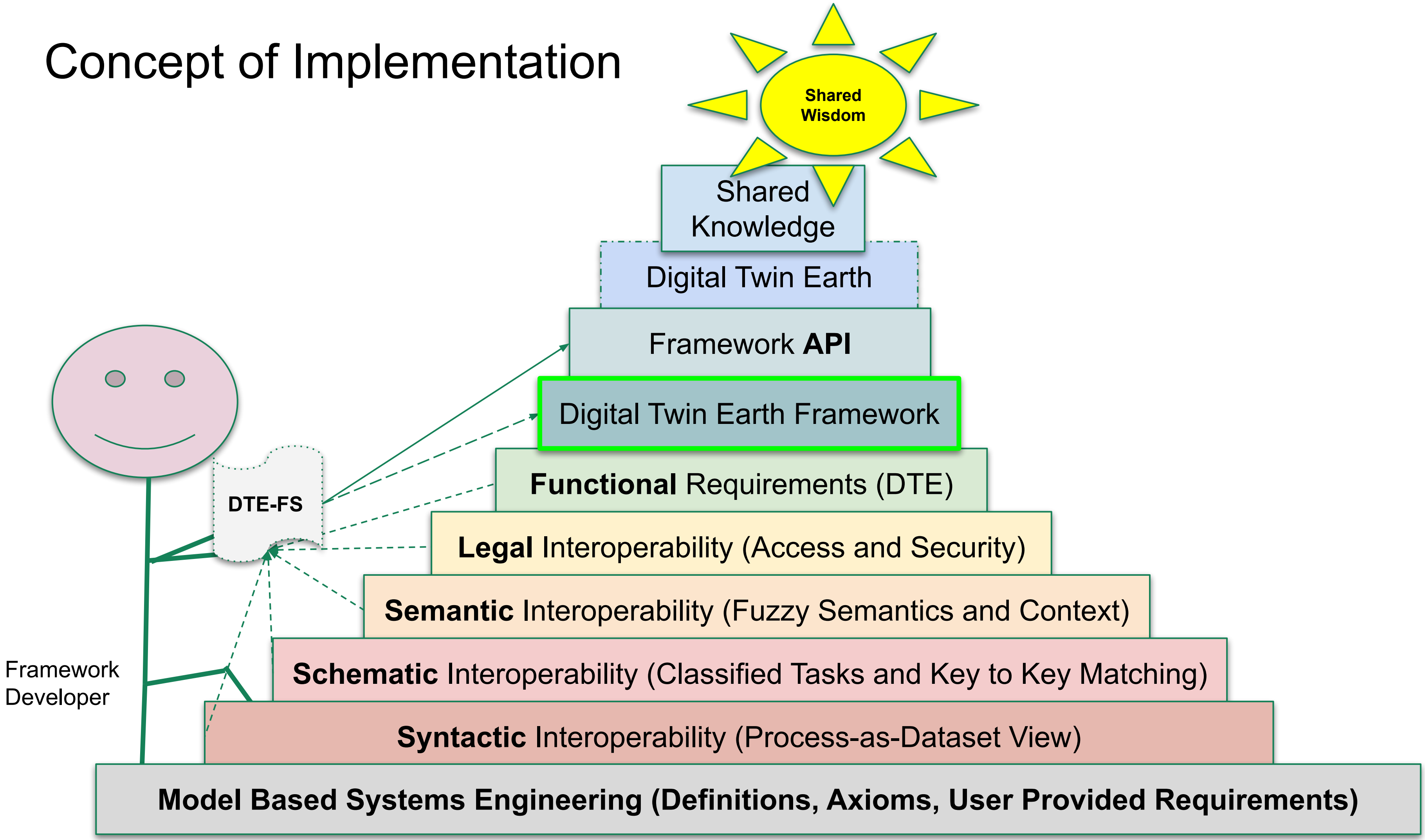


NLP Summit 2022



Deconstructing Analysis-Ready Data

Concept of Implementation



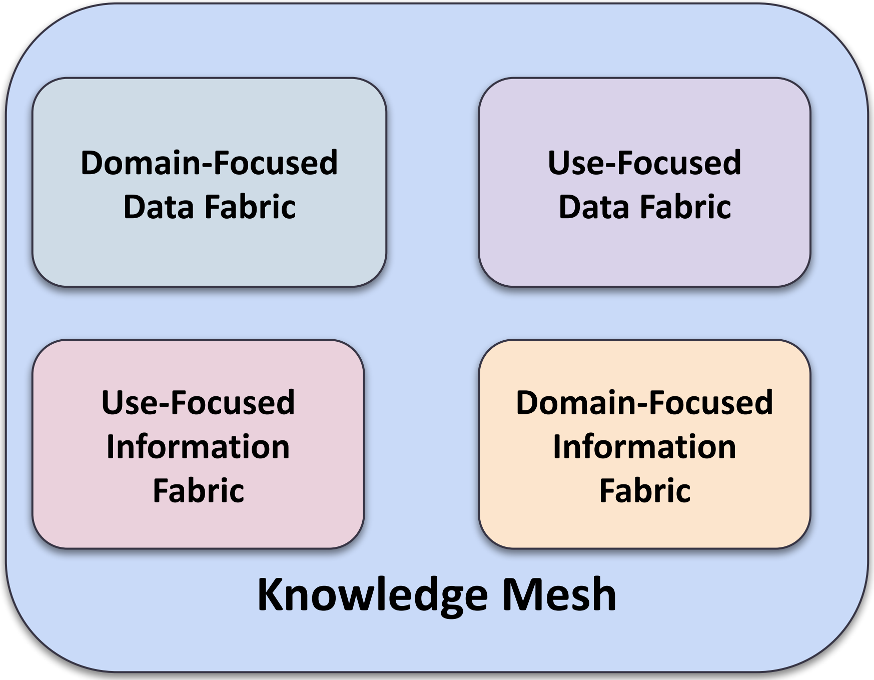
Terminology - Mesh vs Fabric?

Leveraging the Modern Toolbox

It's fuzzy, but very basically,

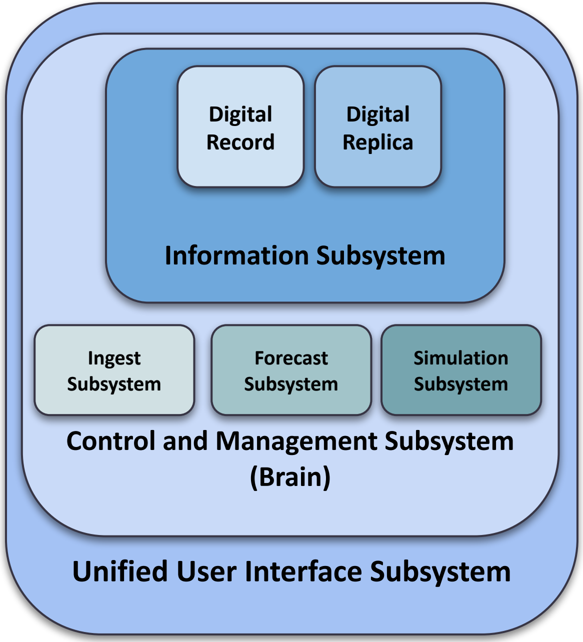
- **Fabric** - centralized governance, focused community, focused access patterns
- **Mesh** - decentralized control, community of communities, integrated access patterns

We need capabilities of both, and are seeing the interplay of these architectural strategies emerge in next-gen motifs.

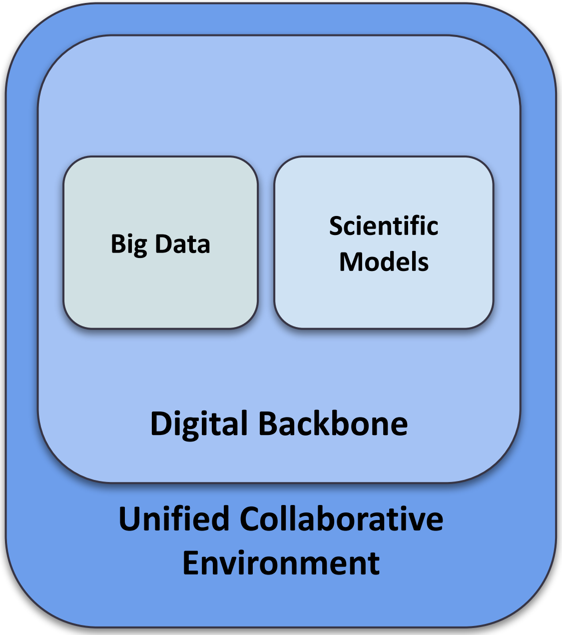


Emerging Motifs in the Next-Gen Landscape

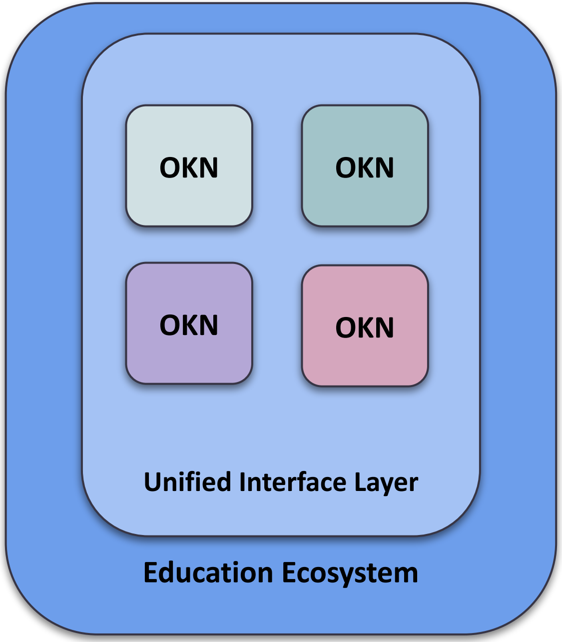
Mapping and aligning our efforts for optimization and interoperability



Earth System Digital Twin
(ESDT)

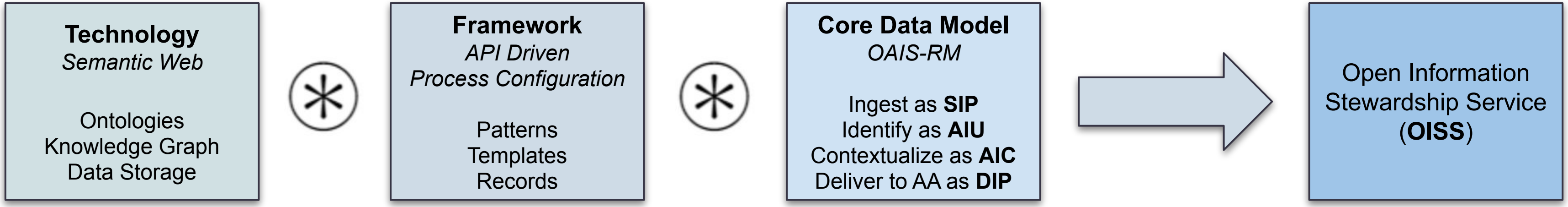


European Digital Twins of the Ocean Ecosystem
(DTO)

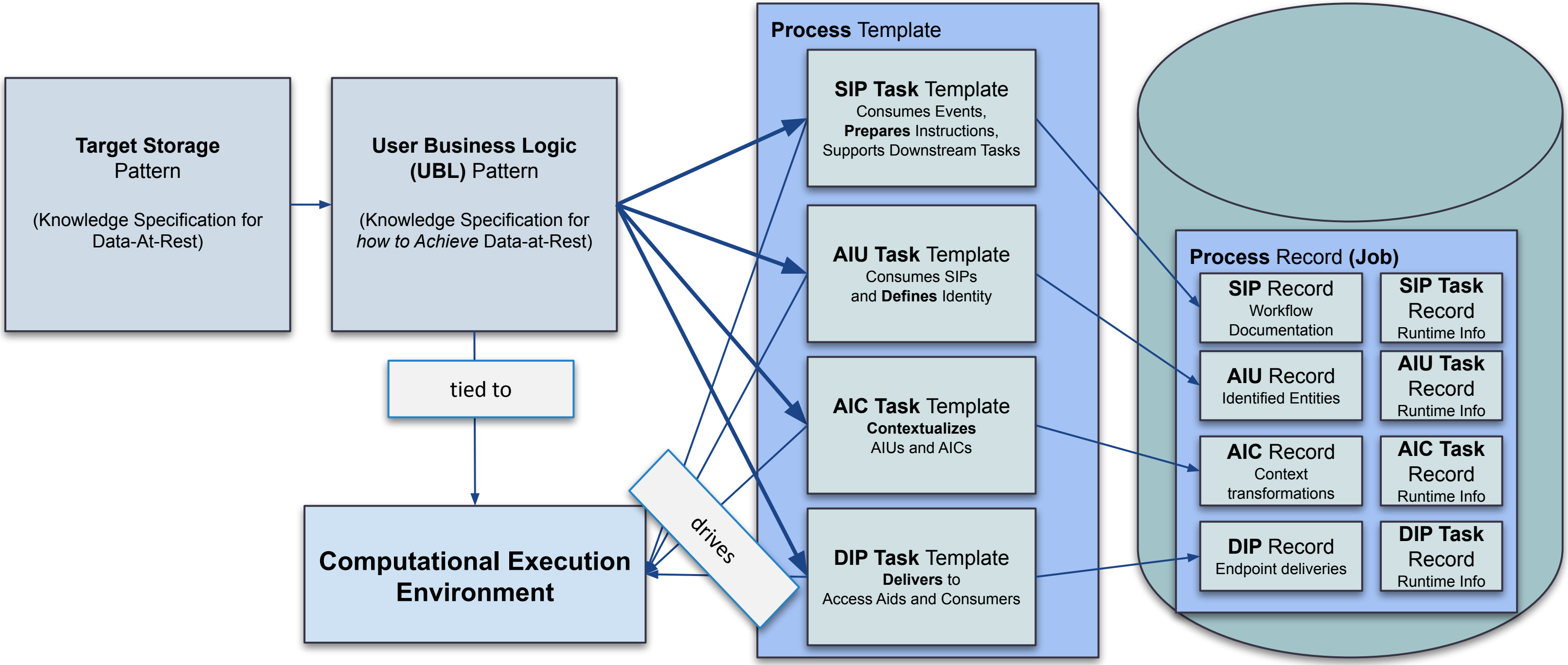


Prototype Open Knowledge Network
(Proto-OKN)

Framework Concept Convolution



Topology

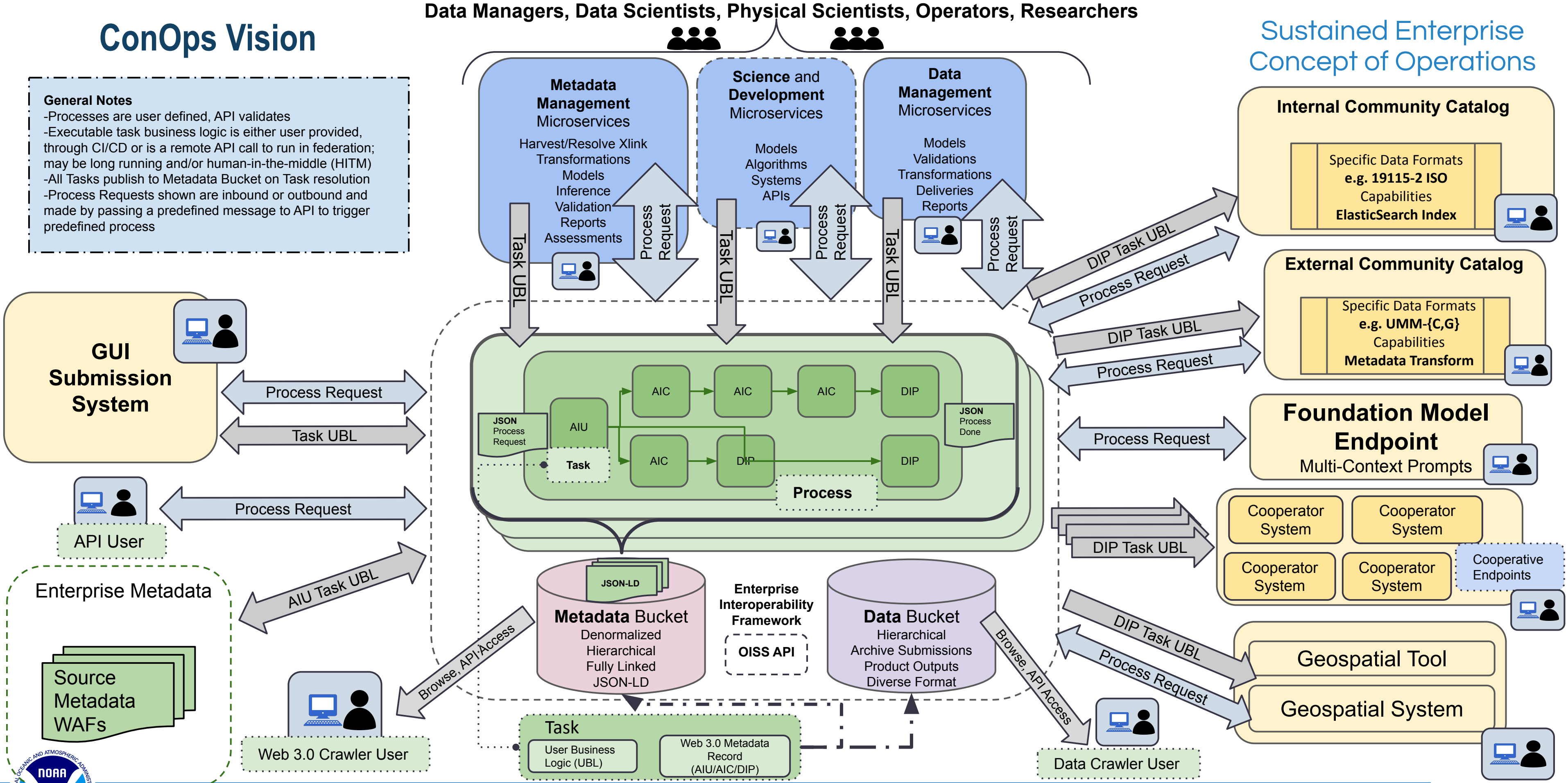


Target Storage Pattern
(Knowledge Specification for Data-At-Rest)
User Business Logic (UBL) Pattern
(Knowledge Specification for how to Achieve Data-at-Rest)
tied to
Computational Execution Environment
drives
Process Template
SIP Task Template
Consumes Events, Prepares Instructions, Supports Downstream Tasks
AIU Task Template
Consumes SIPs and Defines Identity
AIC Task Template
Contextualizes AIUs and AICs
DIP Task Template
Delivers to Access Aids and Consumers
Process Record (Job)
SIP Record
Workflow Documentation
SIP Task Record
Runtime Info
AIU Record
Identified Entities
AIU Task Record
Runtime Info
AIC Record
Context transformations
AIC Task Record
Runtime Info
DIP Record
Endpoint deliveries
DIP Task Record
Runtime Info

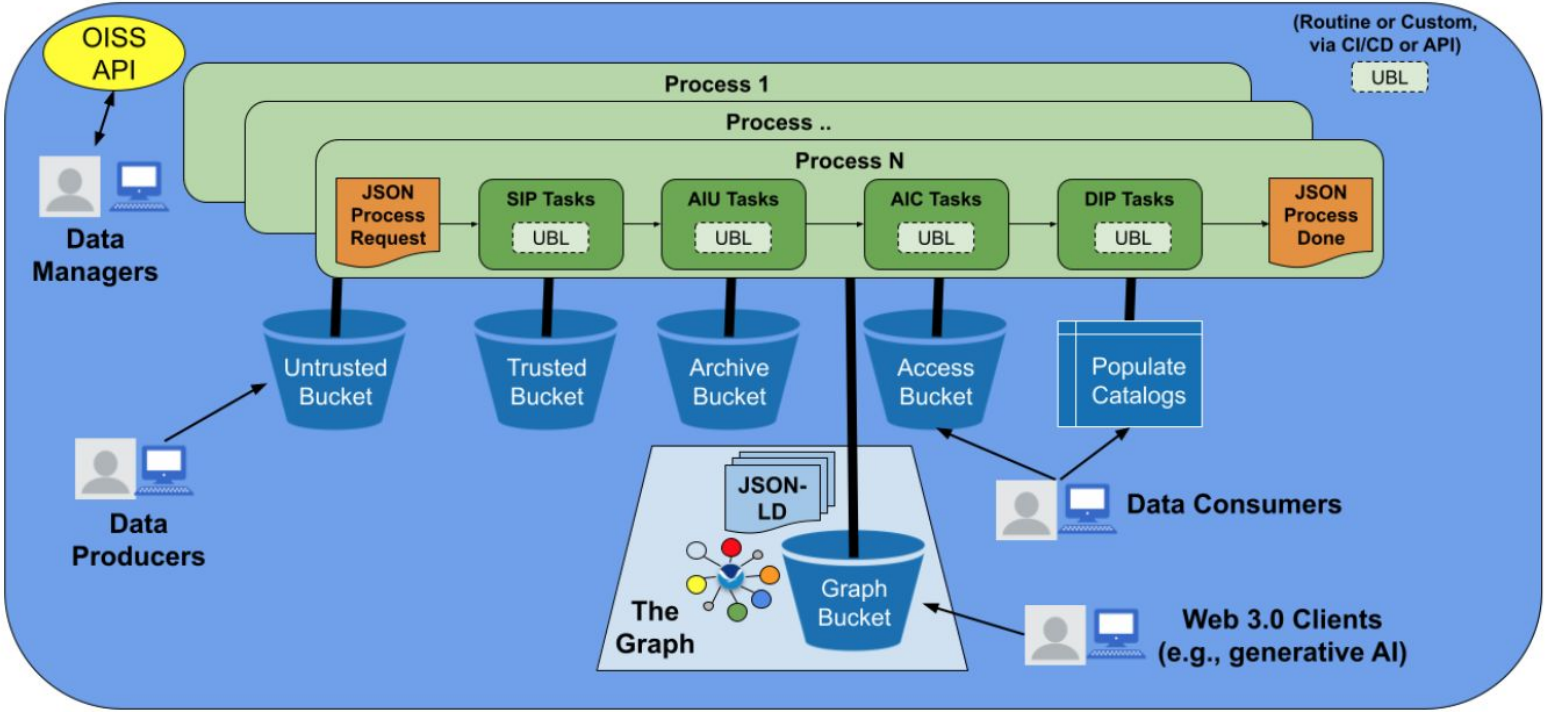
ConOps Vision

Sustained Enterprise Concept of Operations

General Notes
-Processes are user defined, API validates
-Executable task business logic is either user provided, through CI/CD or is a remote API call to run in federation; may be long running and/or human-in-the-middle (HITM)
-All Tasks publish to Metadata Bucket on Task resolution
-Process Requests shown are inbound or outbound and made by passing a predefined message to API to trigger predefined process



ConOps Infrastructure Topology



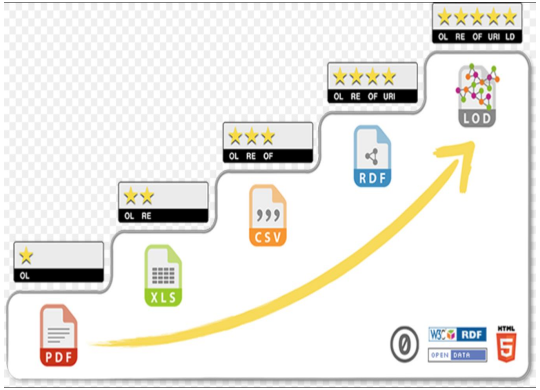
Enabling Iterative Transformation at NOAA

Working toward a sustained cooperative continuous improvement model

Open Information Stewardship Service (OISS) Integration Maturity Scale

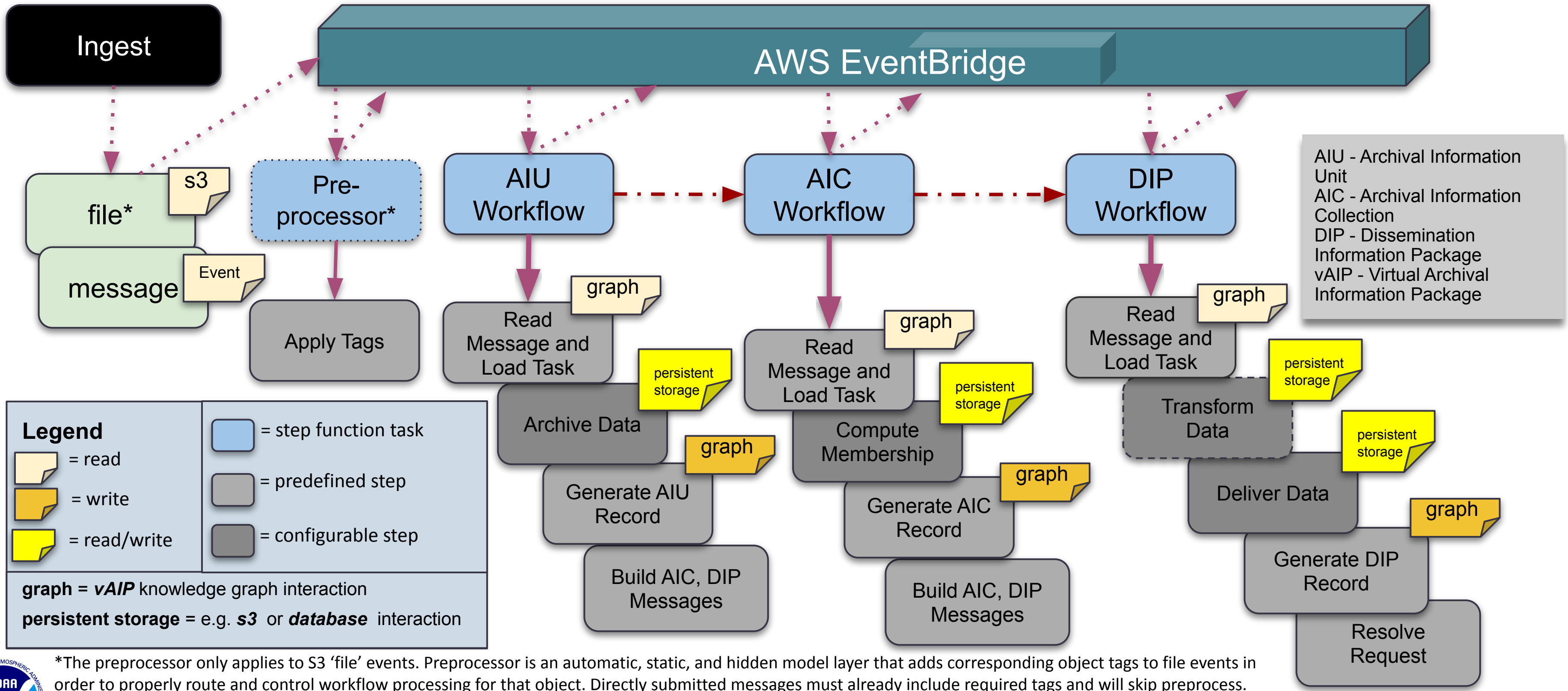
- **Level 4** - Exemplary Integration - Domain Standards Alignment
- **Level 3** - Full Integration - Interoperable Information Models
- **Level 2** - Step Change Integration - Interoperable Resources
- **Level 1** - Baseline Integration - Knowledge of Resources
- **Level 0** - No integration - Baseline NCCF Sandbox Requirements

5 Star Linked Open Data



Information Purview Domain × Recursive Stepwise Integration = Iterative Increases to Openness

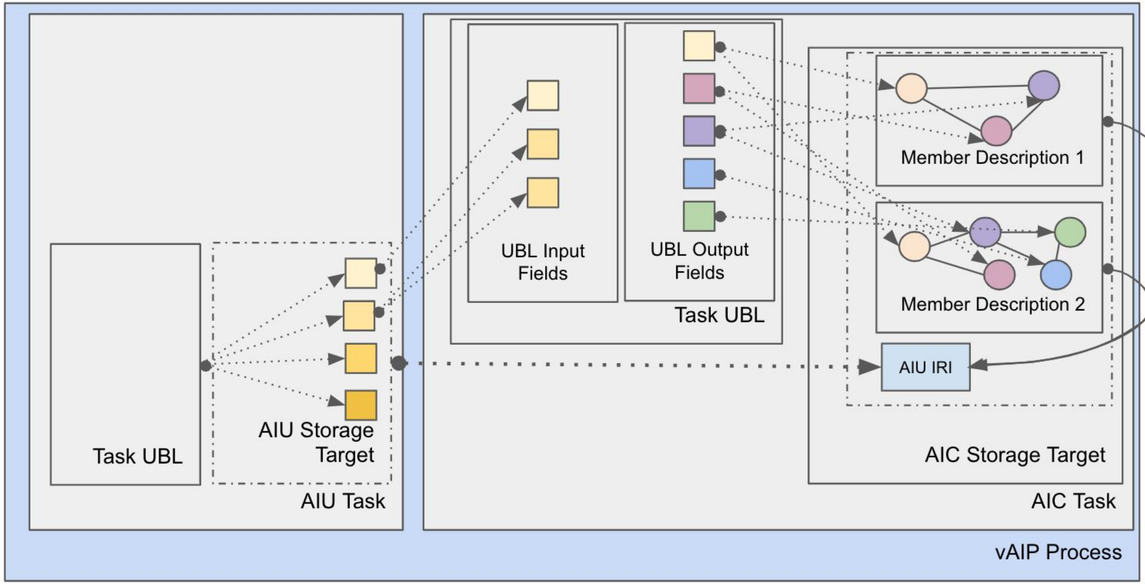
Massive Real-time Concurrency, Small Footprint

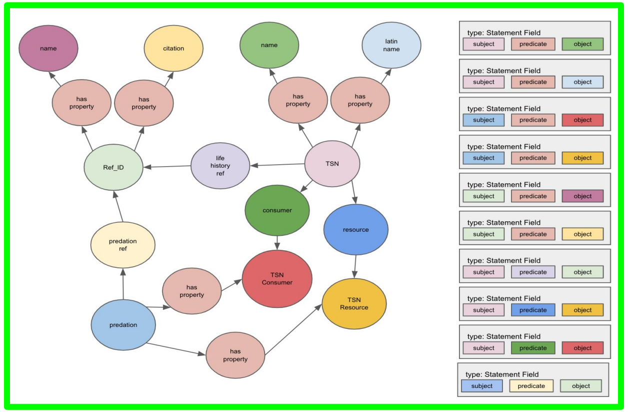


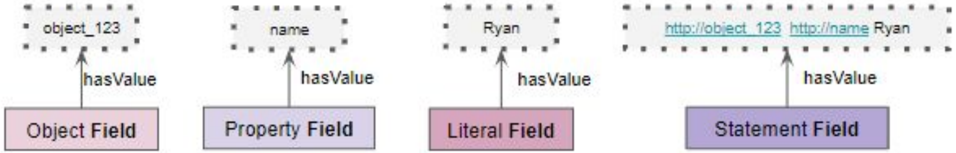
*The preprocessor only applies to S3 'file' events. Preprocessor is an automatic, static, and hidden model layer that adds corresponding object tags to file events in order to properly route and control workflow processing for that object. Directly submitted messages must already include required tags and will skip preprocess.

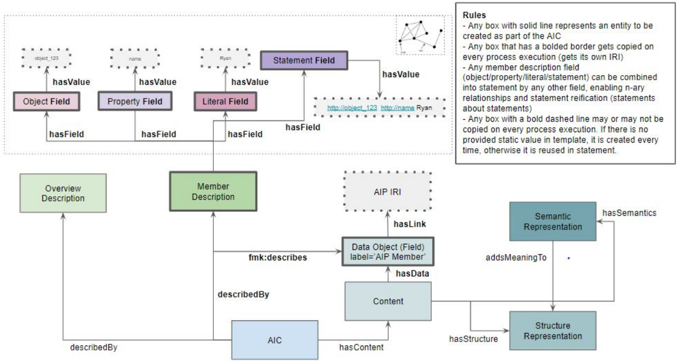
Enormous Amounts of Trusted, High Quality, Fully Contextual, and Regularized Data

Links; texts; documents; input/output parameters for functions; video; audio; biological; ecological; etc.

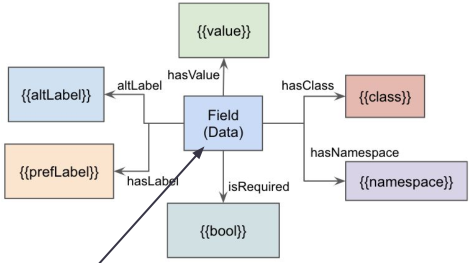


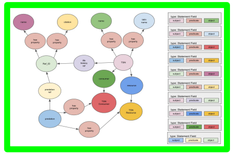


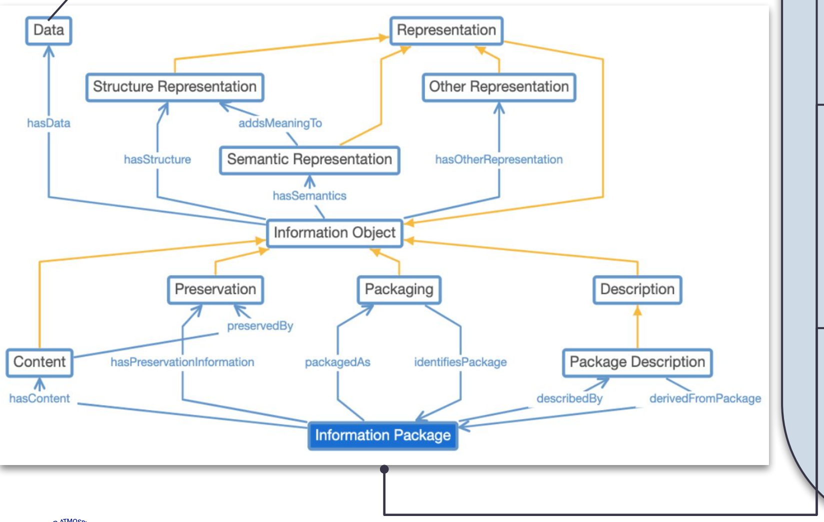




OISS Motifs







AIU
(Entities)

"granule" "weather event" "task" "sensor platform" "species" "hurricane" "estuary" "collection record" "process" "cruise" "wetland study" "data manager" "analysis" "UBL"

AIC
(Representations)

"PII Identified" "Data Quality Score" "Weather Advisory" "Lightning density" "DSRR" "Format Transformation" "Climate Impacts" "Extracted Entities" "Food Web" "Trust Assessment" "Hurricane Path" "Hazard Alert" "User Sentiment" "Similar Datasets" "Crop Yield Prediction" "Similar Orders" "Security Score"

DIP
(Access Disseminations)

"Publish to OIH/ODIS" "Publish to data.gov" "Broadcast Text Message" "Notify Internet of Water" "Publish to One Stop" "Return API Response" "Send Email Alert" "Publish to DPMF" "Publish to CMR" "Feed foundation model"

Pattern Flavors (*SHACL Validated*)

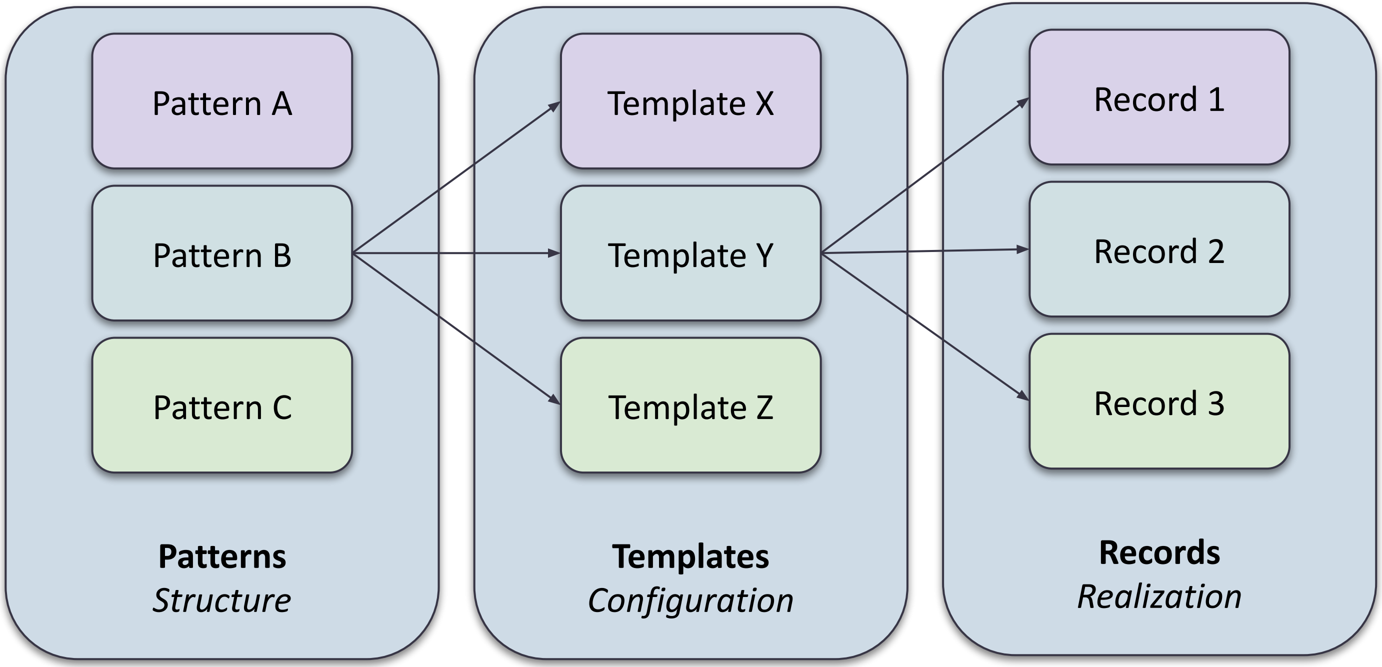
Motif Search Spaces

Motifs (subgraphs) are enablers for:
-**Stable API interfaces**
-**Optimized searches**
-**Feature models**
-**Visualizations**

All OISS **explicit motifs** are **OAIS Information Packages** Reference model defined user patterns.

Each explicit motif has different SHACL characteristics - for example, **AICs support n-ary relationships** in their member descriptions.

These are the motifs that are defined and known - there are other motifs emergent in the system data due to its linked data nature



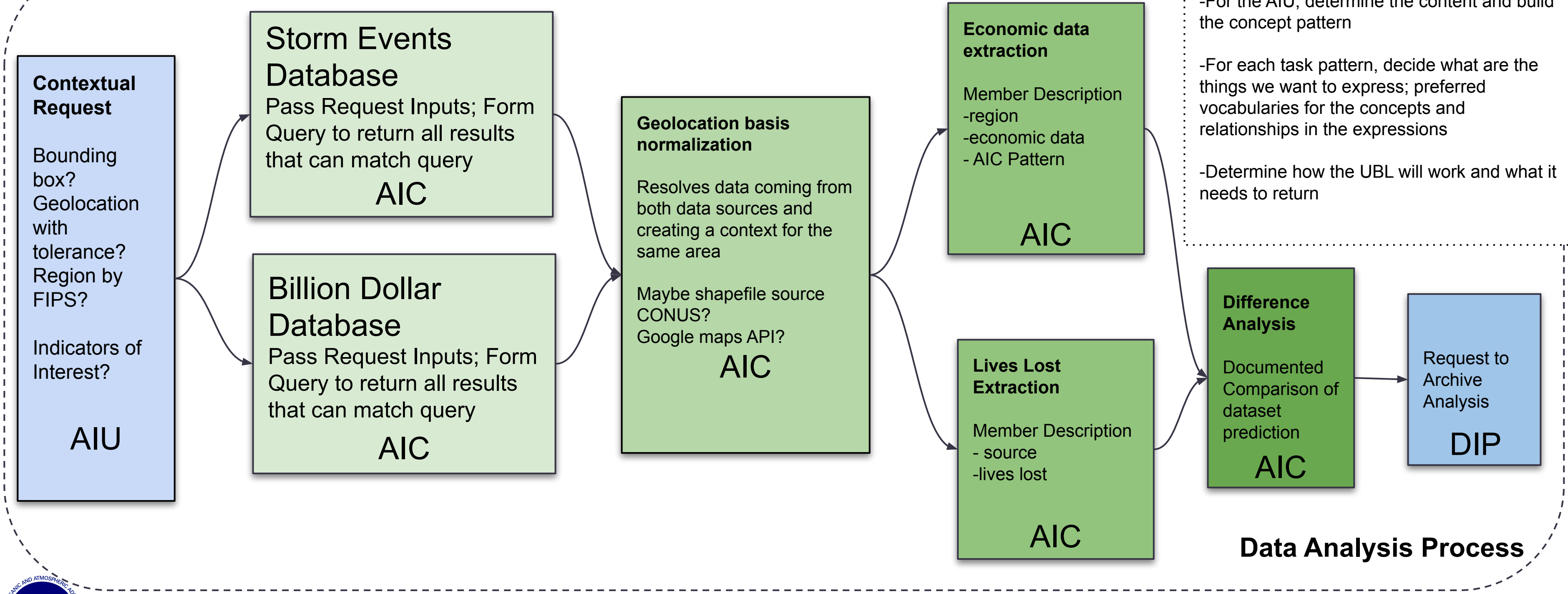
OAR-NCEI Societal Data Partnership

Questions to Ask

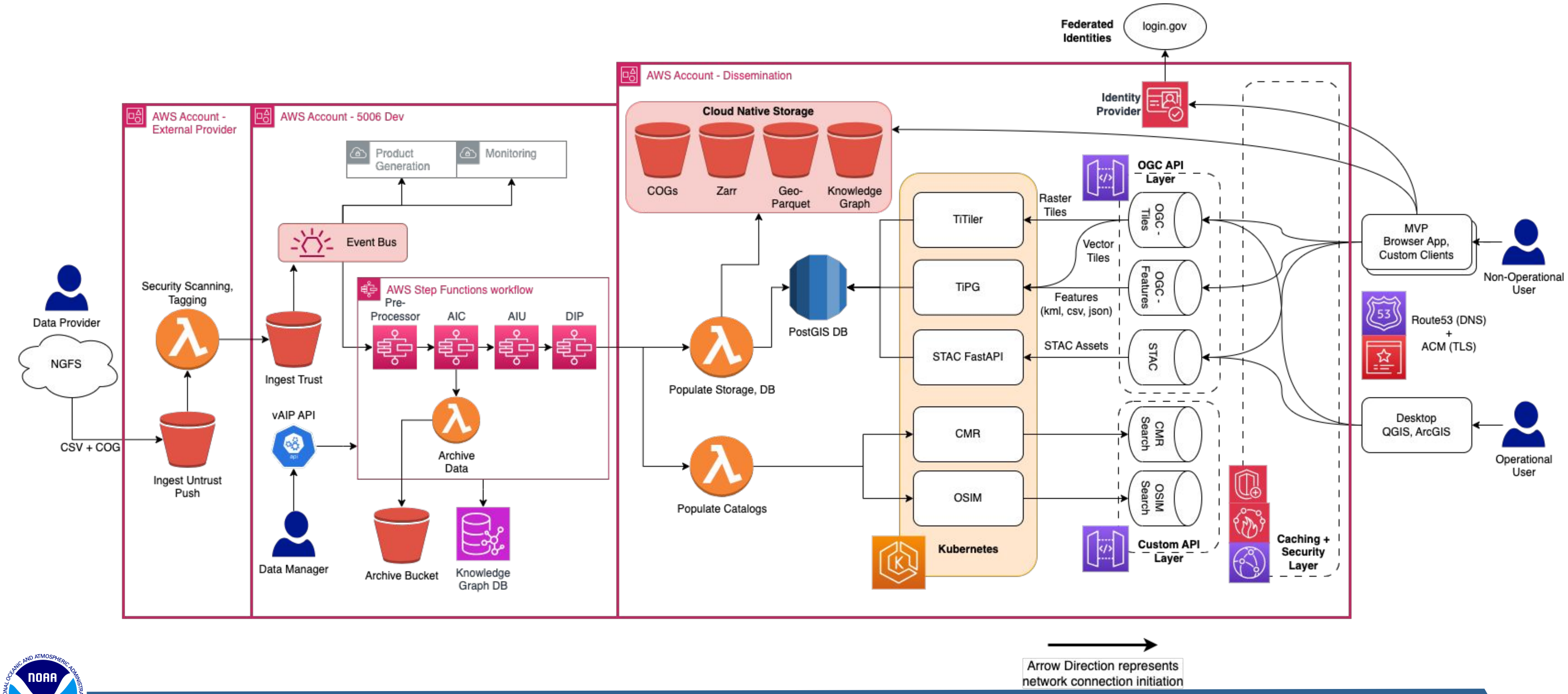
-For the AIU, determine the content and build the concept pattern

-For each task pattern, decide what are the things we want to express; preferred vocabularies for the concepts and relationships in the expressions

-Determine how the UBL will work and what it needs to return

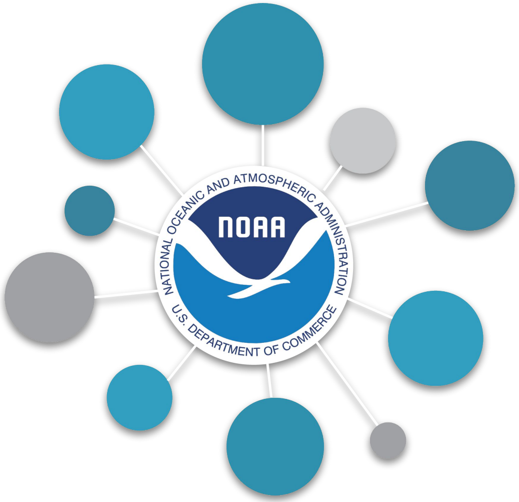


Multi-transform and Multi-access for Wildland Fire Storefront



Current Activities

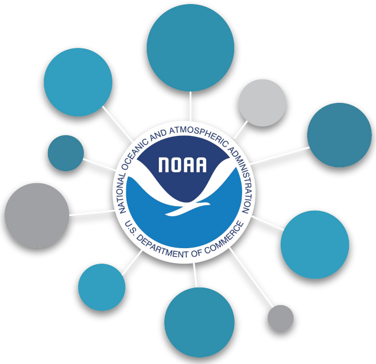
- **Seeding the mesh** with early adopter processes and resources
  - Lots of data streams and system integrations being onboarded or planned - Nexrad, CCOR, OISST, Faraday Cup, Extended Continental Shelf, Passive Acoustic Data, IGRA, World Ocean Database, etc.
- Developing and delivering **educational materials** for mesh adoption and maturation
  - Jupyter notebooks; trainings; ‘teaching to fish’ with the API and using the JSON-LD files and various data formats that are produced
- Identifying **beneficial partnerships** and **integration pathways**
  - E.g. across NOAA LOs; with NASA AIST and CNES in Digital Twin contexts; with NSF in OKN contexts; with OIH/ODIS in specific Ocean datastream contexts
- Developing **access processes** using **advanced AI/ML**
  - E.g. use of Retrieval Augmented Generation (RAG) to build out **conversational discovery** and access Processes; automated metadata normalization
- Building out the **marketplace**
  - Working with users to build out system resources for sharing



Backups

Mesh Highlight Summary

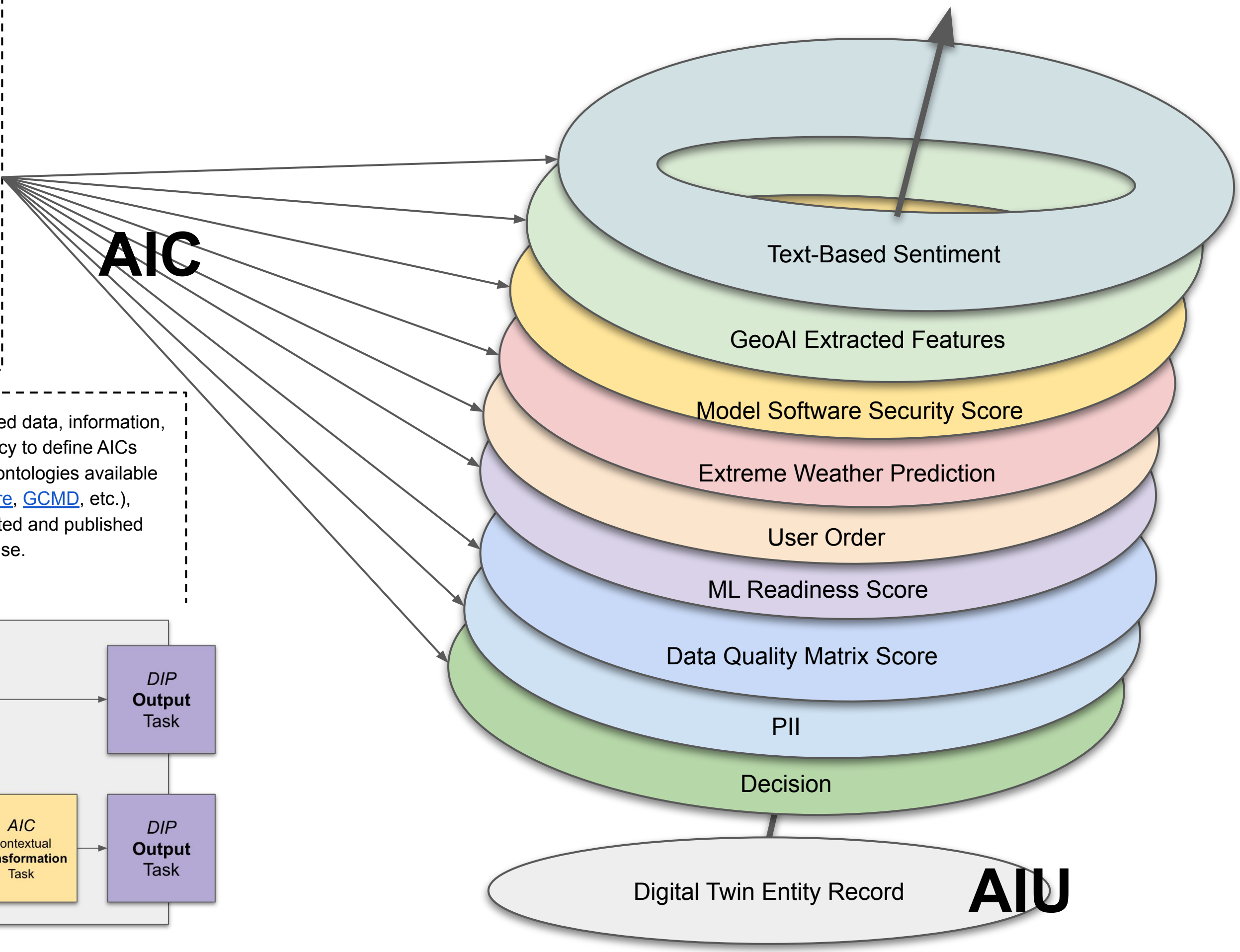
- Provides a full capability **metadata layer as** an **enterprise orchestration** service
  - Supports all **Digital Twin Earth Framework requirements** as specified by the DTE-FS
  - Maps as an AIST ESDT Architecture
  - Satisfies NARA requirements, intrinsic and largely automated support for PARR 2.0, FAIR, CARE, TRUST
  - As **cheap** as possible - cost optimized **at full capability**
  - As **flexible** as possible - any process, entity, expression, or access pattern; any user code in any compute context; agile reference architecture for re-implementation
    - Allows evolutionary growth in all exposed aspects
  - As **simple** as possible - no message adapters, pip install, small purposeful API
- Emphasizes **facts** and **traceability**
  - Everything is data; process is data; full provenance at the task level published in real-time on every process event
- Enables **definition** and **contextualization** of **anything** (entities, transformations, disseminations)
  - Inbound and outbound processes are the same; capturing use information; semantic records of function call inputs and outputs for model training; only running on request; API definition of everything
  - Provides a Motif basis for everything in larger connected context - **Motifs are critical** for enabling analyses
- Enables easy and automatable **intra-agency** and **inter-agency** interoperability
  - Process lets us connect with (and trace) federated APIs; supports interoperability with legal and regulatory requirements; private data
  - Open participation
- Supports **exploitation by Next-Gen capabilities** - multimodal foundation models, virtual reality environments, ‘Spatial Computing’
- Guards against future migrations (**assumes eventual sunset up front**) - everything is data, no ‘code-lock’

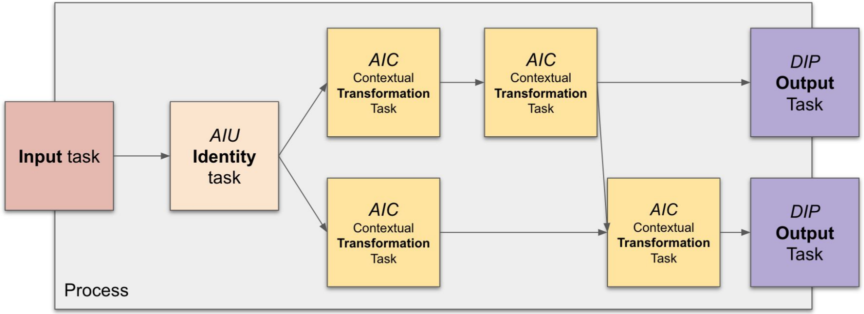


Linked Independent Information Models

- Structured KOS
- FAIR
  - Findable
  - Accessible
  - Interoperable
  - Reusable
- CARE
  - Fine-grain access rights
  - Legal Interoperability
- **Knowledge Graph** held
- Independent spaces
- API Discovery/Document Browsing

The **AIC** provides controlled access to the ‘full graph’ of semantically encoded data, information, and knowledge, allowing users authorized by associated Access Rights policy to define AICs that implement Patterns using 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 that are constructed and published automatically on AIC Pattern deployment for system-wide discovery and reuse.

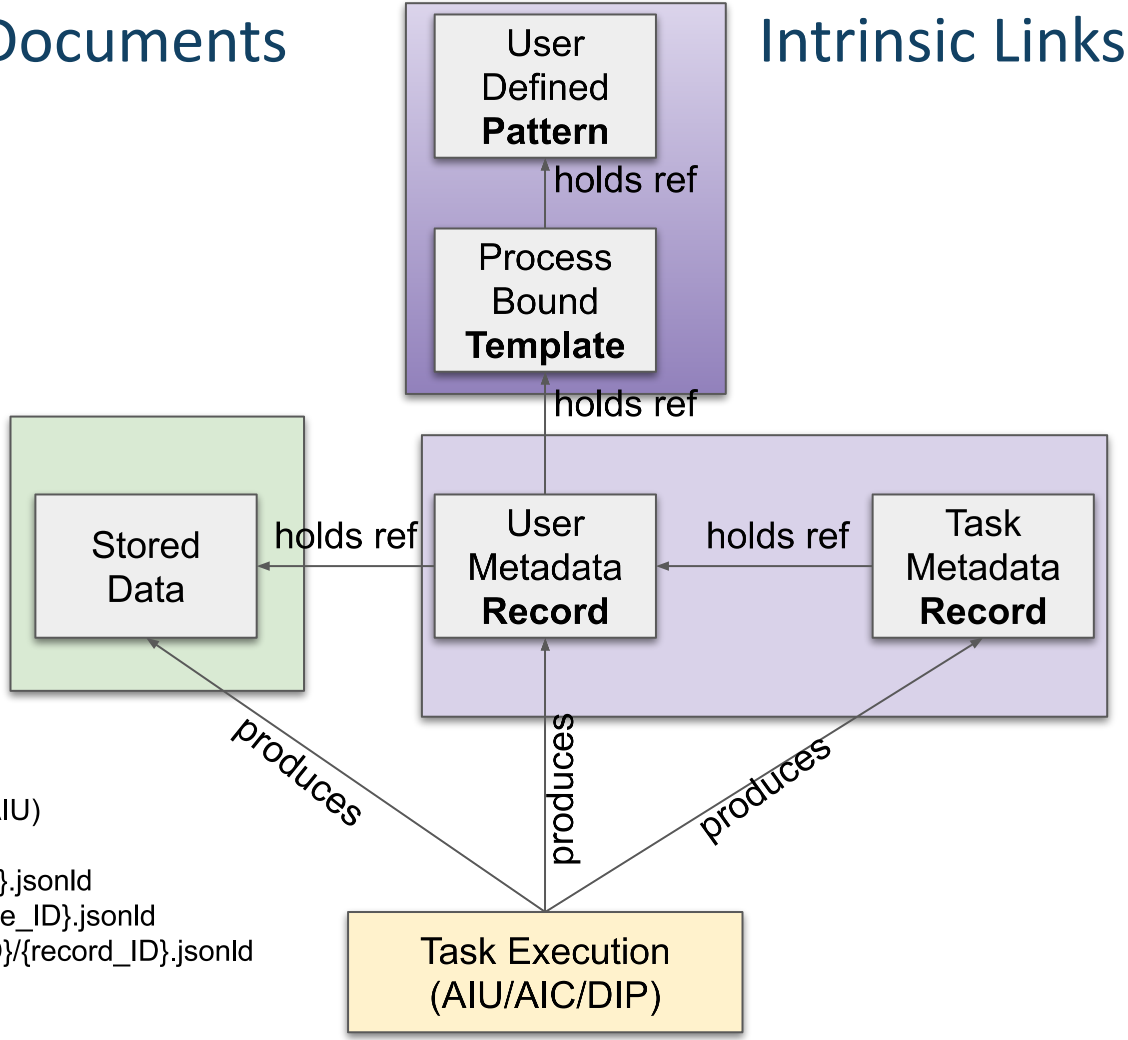




AIU/AIC/DIP Metadata Documents

Intrinsic Links

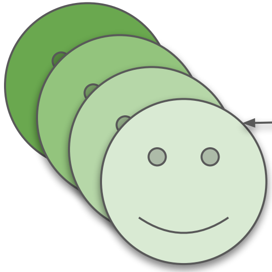
- S3 Archive/Access Bucket
- S3 Metadata Bucket
- S3 Metadata Bucket/Neptune/Cache
- Step Function



Metadata Bucket Hierarchy (example layout for AIU)

s3://patterns/storage/aiu/{pattern_ID}/{pattern_ID}.jsonld
s3://templates/storage/aiu/{template_ID}/{template_ID}.jsonld
s3://records/storage/aiu/{template_ID}/{record_ID}/{record_ID}.jsonld

Direct, Immediate Metadata Access



Static Page 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

Creating a Pattern

```
pattern = ontology.IdentityStoragePattern(session, "Granule")

reference = pattern.add_reference("Archive PID", field={'title': 'Value', 'is_link': False})
semantic = reference.add_semantic_representation("Description")
structure = reference.add_structure_representation("Format")
semantic.node.addsMeaningTo.append(structure.node)

checksum = pattern.add_fixity("Checksum", field={'title': 'Value', 'is_link': False})
checksum.add_extra_representation("Purpose")
checksum_semantic = checksum.add_semantic_representation("Definition")
checksum_structure = checksum.add_structure_representation("Format")
checksum_semantic.node.addsMeaningTo.append(checksum_structure.node)

file_content = pattern.set_content("File", field={'is_link': True})
file_semantic = file_content.add_semantic_representation("File Schema")
file_structure = file_content.add_structure_representation("File Format")
file_semantic.node.addsMeaningTo.append(file_structure.node)
file_content.add_extra_representation("Decoding Software")
file_content.add_extra_representation("Visualization Software")

desc = pattern.add_description("Product Description")
desc.add_semantic_representation("Product Description Semantic Rep")
desc.add_structure_representation("Product Description Structure Rep")


ar1 = pattern.add_access_rights("Distribution License")
ar1.add_semantic_representation("Distribution License Semantic Rep")
ar1.add_structure_representation("Distribution License Structure Rep")
```

Create a new pattern (here, AIU)

Add AIP-required information objects^

- Content
- Descriptive Information
- Packaging
- PDI
  - Reference
  - Fixity
  - Access Rights
  - Context*
  - Provenance*

Add representation IOs to each IO

- structure
- semantic
- 'extra' (other)

Notes - Can add any number of Fields (DOs) per IO
Fields are semantic (classed and namespaced) links or values

*Automatically satisfied, but user can extend
^Can add greater or equal to the Information Package Subtype (AIU/AIC/DIP) Requirements

Creating a Process (Pattern to Workflow)

```
process = ProcessTemplate(session, purpose="E2E Granule Process", utility="Test process template")
process_namespace = process.namespace

## Initialize the IdentityStorageTemplate using a premade pattern
granule_pattern = granule(session)
aiu_template = IdentityStorageTemplate(session, source_pattern=granule_pattern)

task = IdentityTaskTemplate(session, target=aiu_template, purpose="E2E Granule AIU Task", utility="Test task tem
process.set_aiu_task(task)
ubl = UBLTemplate(session, namespace=task.namespace)
task.set_ubl(ubl)

# This sets the AWS ARN for the step function executing this task (the aiu-workflow step function)
task.set_packaging_field(packaging="Resource Identifier", field="Value", value="arn:aws:states:us-east-1:5602713

## Set the AWS ARN for the UBL step function that gets executed in the AIU Task (archive-data)
## This is required in order for the AIU Worflow to call its UBL
ubl.set_packaging_field(packaging="Resource Identifier", field="Value", value="arn:aws:states:us-east-1:56027137

## Now we add all the input fields to the UBL
# The inputs below here are all "hardcoded" at the UBLTemplate level
archive_file_lambda_arn = ubl.add_input("Archive File ARN", value="arn:aws:lambda:us-east-1:560271376700:functio
create_uuid_lambda_arn = ubl.add_input("Create UUID ARN", is_required=False, value="arn:aws:lambda:us-east-1:560
destination_prefix = ubl.add_input("Destination Prefix", value="copy")
destination_key = ubl.add_input("Destination Key", is_required=False, value="")
destination_bucket = ubl.add_input("Destination Bucket", value="nccf-dev-archive-us-east-1-560271376700")
algorithm = ubl.add_input("Checksum Algorithm", value="MD5")
retention = ubl.add_input("Retention Time", value=20)
ubl.add_input("Archive Status", is_required=False, value="")
```

Create a Process (template)

Add Tasks (AIU/AIC/DIP)

For each Task, specify what we're storing (Pattern), and how to produce it (UBL - user business logic)

For UBL, specify the 'field map' - what goes in, what comes out

For a Task Pattern, optionally specify static values

Connect Task Internally (UBL key to key matching)

Connect Tasks externally (Pattern to UBL matching

Dynamic Resource Analysis During Creation


```xml
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
         xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
         xmlns:owl="http://www.w3.org/2002/07/owl#"
         xml:base="https://ncei.noaa.gov/vaip/pattern/storage/aiu/6b587130-829e
         xmlns="https://ncei.noaa.gov/vaip/pattern/storage/aiu/6b587130-829e-42
         xmlns:core="https://ncei.noaa.gov/vaip/core#"
         xmlns:core2="http://www.w3.org/2004/02/skos/core#">

<owl:Ontology rdf:about="https://ncei.noaa.gov/vaip/pattern/storage/aiu/6b58713

<core:ArchivalInformationUnit rdf:about="https://ncei.noaa.gov/vaip/pattern/sto
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#NamedIndividual"/>
  <core:hasContentInformation rdf:resource="https://ncei.noaa.gov/vaip/pattern/
  <core2:prefLabel rdf:datatype="http://www.w3.org/2001/XMLSchema#string">weath
</core:ArchivalInformationUnit>

<core:ContentInformationObject rdf:about="https://ncei.noaa.gov/vaip/pattern/st
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#NamedIndividual"/>
  <core:hasDataObject rdf:resource="https://ncei.noaa.gov/vaip/pattern/storage/
  <core:hasDataObject rdf:resource="https://ncei.noaa.gov/vaip/pattern/storage/
  <core2:prefLabel rdf:datatype="http://www.w3.org/2001/XMLSchema#string">extre
  <core2:altLabel rdf:datatype="http://www.w3.org/2001/XMLSchema#string">foreca
  <core2:altLabel rdf:datatype="http://www.w3.org/2001/XMLSchema#string">model<
  <core2:altLabel rdf:datatype="http://www.w3.org/2001/XMLSchema#string">data d
</core:ContentInformationObject>

<core:DigitalObject rdf:about="https://ncei.noaa.gov/vaip/pattern/storage/aiu/6
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#NamedIndividual"/>
  <core2:prefLabel rdf:datatype="http://www.w3.org/2001/XMLSchema#string">event
  <core:hasLink rdf:datatype="http://www.w3.org/2001/XMLSchema#string">{{d2dfab
</core:DigitalObject>

<core:PhysicalObject rdf:about="https://ncei.noaa.gov/vaip/pattern/storage/aiu/
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#NamedIndividual"/>
  <core2:prefLabel rdf:datatype="http://www.w3.org/2001/XMLSchema#string">event
  <core:hasValue rdf:datatype="http://www.w3.org/2001/XMLSchema#string">{{e0f98
</core:PhysicalObject>
```


resource.serialize()

```python
print(granule.example(ancestors=True))
```

✓ 0.0s

```
Archival Information Unit(ancestor 0): An example of an AIU might be a granule; a file containi

Archival Information Package(ancestor 1): An example of an archival information package might b

Information Package(ancestor 2): An information package might be a SIP that represents a proces
```

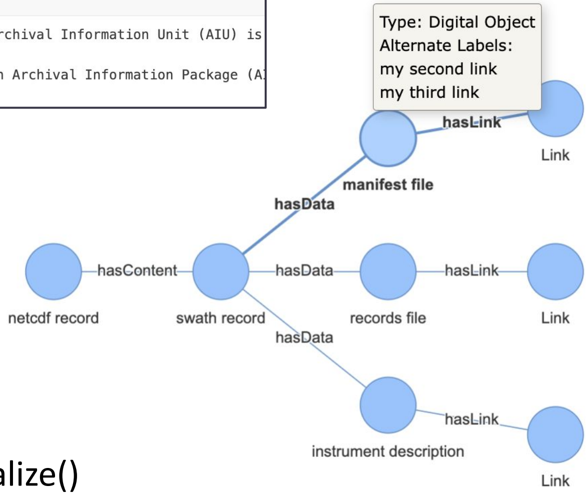
resource.example()

```python
print(granule.definition(ancestors=True, limit=1))
```

✓ 0.0s

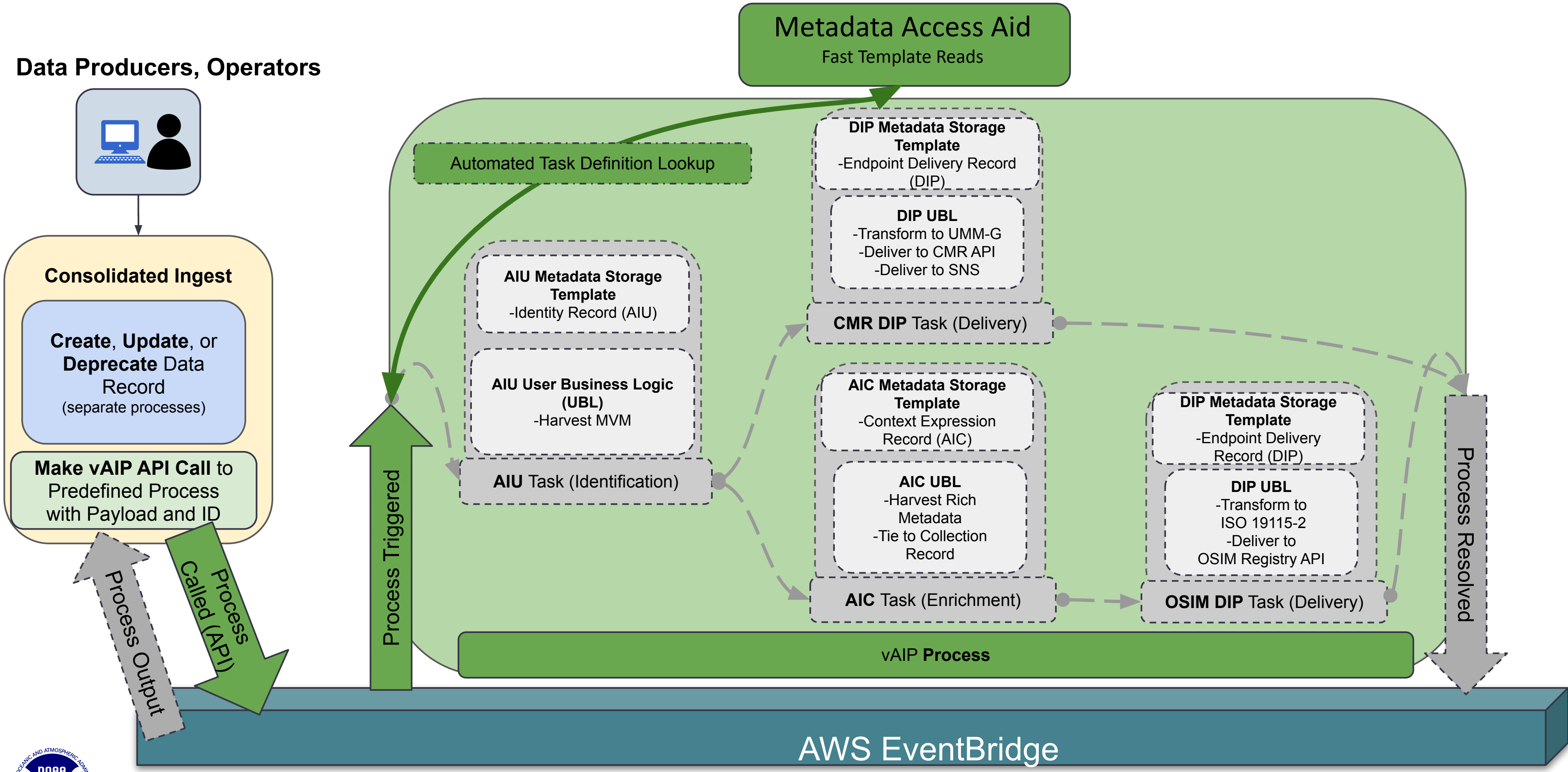
```
Archival Information Unit(ancestor 0): An Archival Information Unit (AIU) is

Archival Information Package(ancestor 1): An Archival Information Package (A
```

resource.definition()

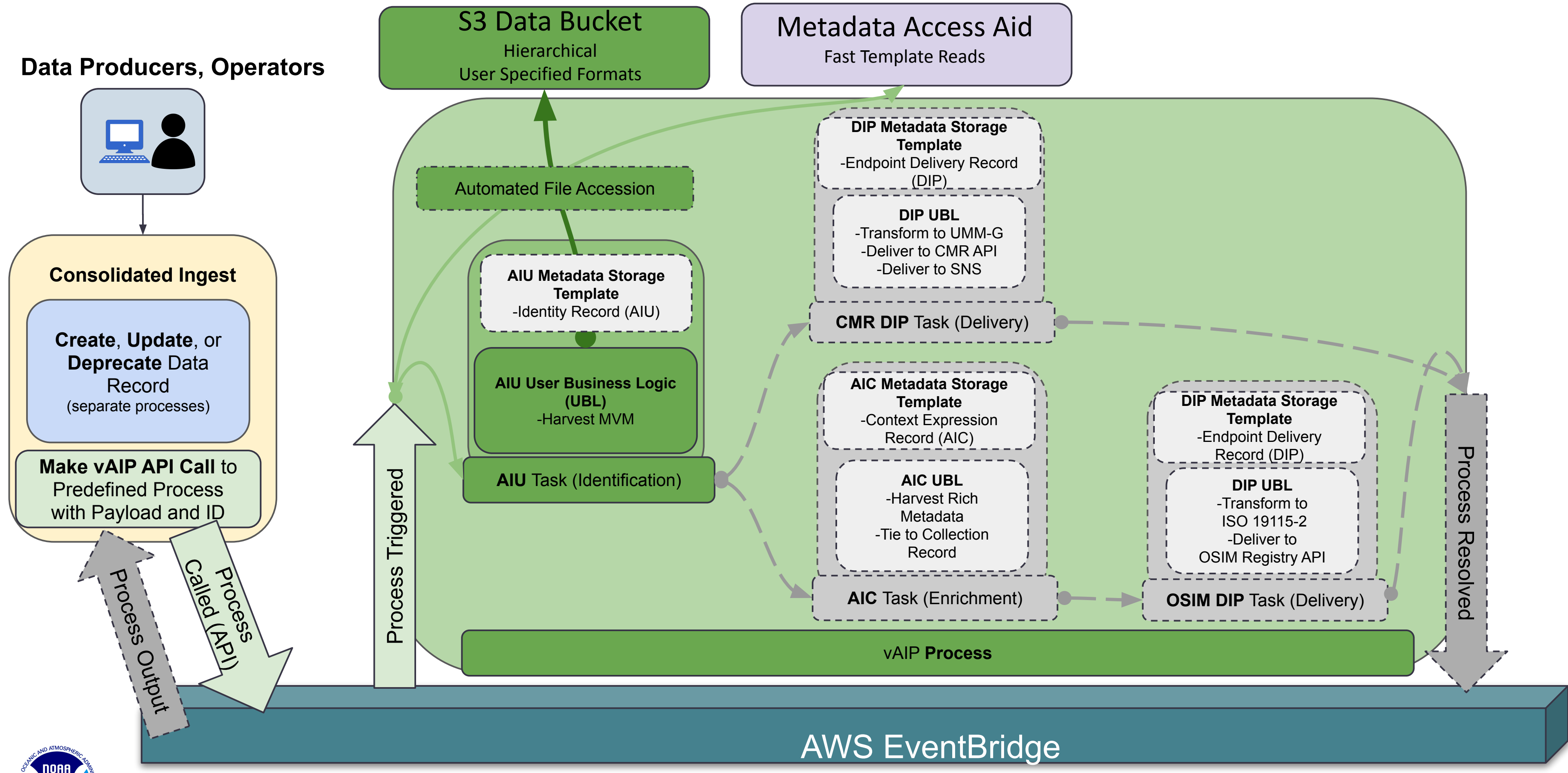


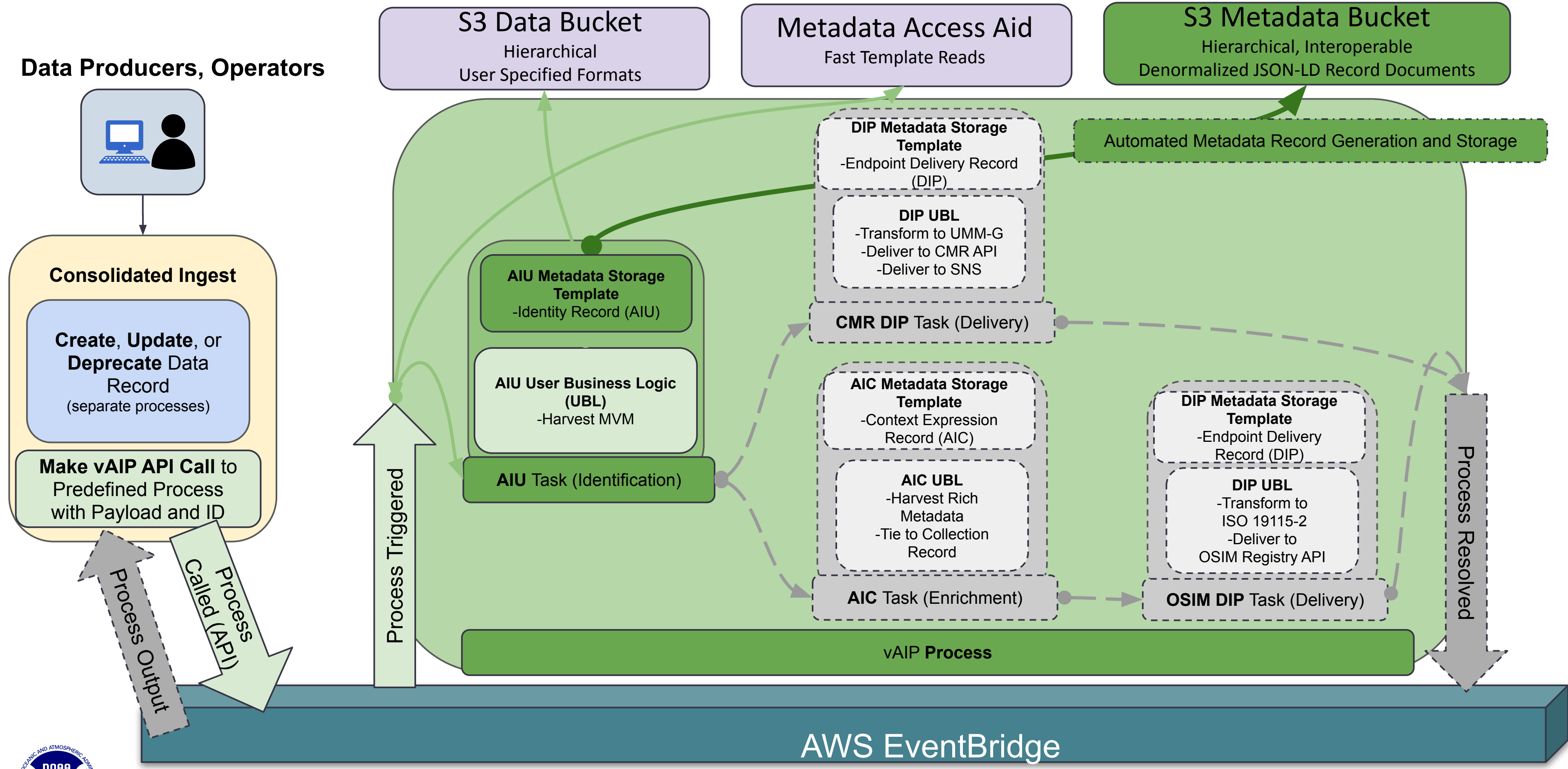
resource.visualize()

note: expanded and improved analytic capability is on the backlog as a higher order API library

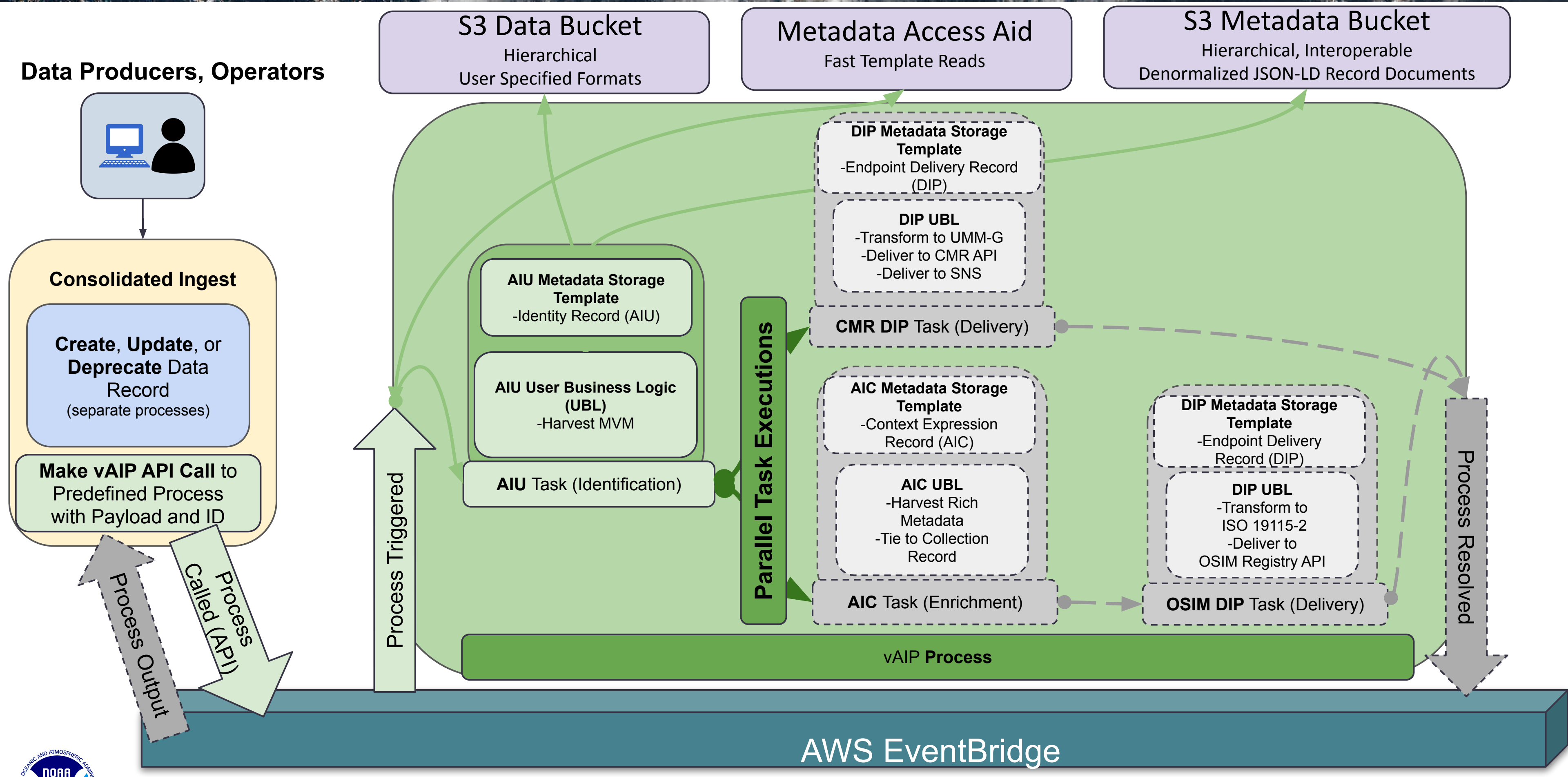


Metadata Access Aid
Fast Template Reads
Data Producers, Operators
Consolidated Ingest
Create, Update, or Deprecate Data Record (separate processes)
Make vAIP API Call to Predefined Process with Payload and ID
Process Output
Process Called (API)
Process Triggered
Automated Task Definition Lookup
AIU Metadata Storage Template
-Identity Record (AIU)
AIU User Business Logic (UBL)
-Harvest MVM
AIU Task (Identification)
DIP Metadata Storage Template
-Endpoint Delivery Record (DIP)
DIP UBL
-Transform to UMM-G
-Deliver to CMR API
-Deliver to SNS
CMR DIP Task (Delivery)
AIC Metadata Storage Template
-Context Expression Record (AIC)
AIC UBL
-Harvest Rich Metadata
-Tie to Collection Record
AIC Task (Enrichment)
DIP Metadata Storage Template
-Endpoint Delivery Record (DIP)
DIP UBL
-Transform to ISO 19115-2
-Deliver to OSIM Registry API
OSIM DIP Task (Delivery)
Process Resolved
vAIP Process
AWS EventBridge

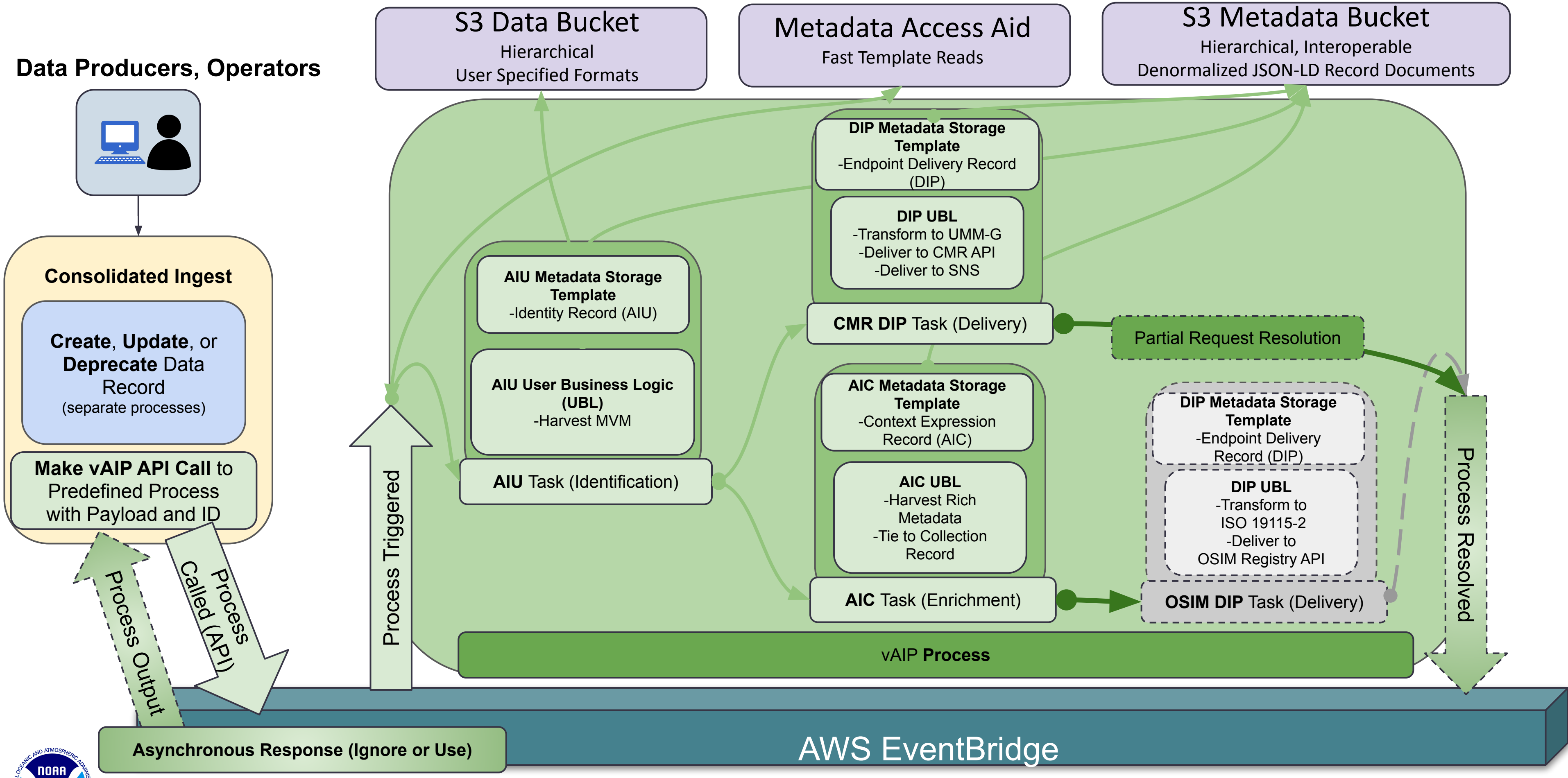

S3 Data Bucket
Hierarchical
User Specified Formats
Metadata Access Aid
Fast Template Reads
Data Producers, Operators
Consolidated Ingest
Create, Update, or Deprecate Data Record
(separate processes)
Make vAIP API Call to Predefined Process with Payload and ID
Process Output
Process Called (API)
Process Triggered
Automated File Accession
AIU Metadata Storage Template
-Identity Record (AIU)
AIU User Business Logic (UBL)
-Harvest MVM
AIU Task (Identification)
DIP Metadata Storage Template
-Endpoint Delivery Record (DIP)
DIP UBL
-Transform to UMM-G
-Deliver to CMR API
-Deliver to SNS
CMR DIP Task (Delivery)
AIC Metadata Storage Template
-Context Expression Record (AIC)
AIC UBL
-Harvest Rich Metadata
-Tie to Collection Record
AIC Task (Enrichment)
DIP Metadata Storage Template
-Endpoint Delivery Record (DIP)
DIP UBL
-Transform to ISO 19115-2
-Deliver to OSIM Registry API
OSIM DIP Task (Delivery)
Process Resolved
vAIP Process
AWS EventBridge

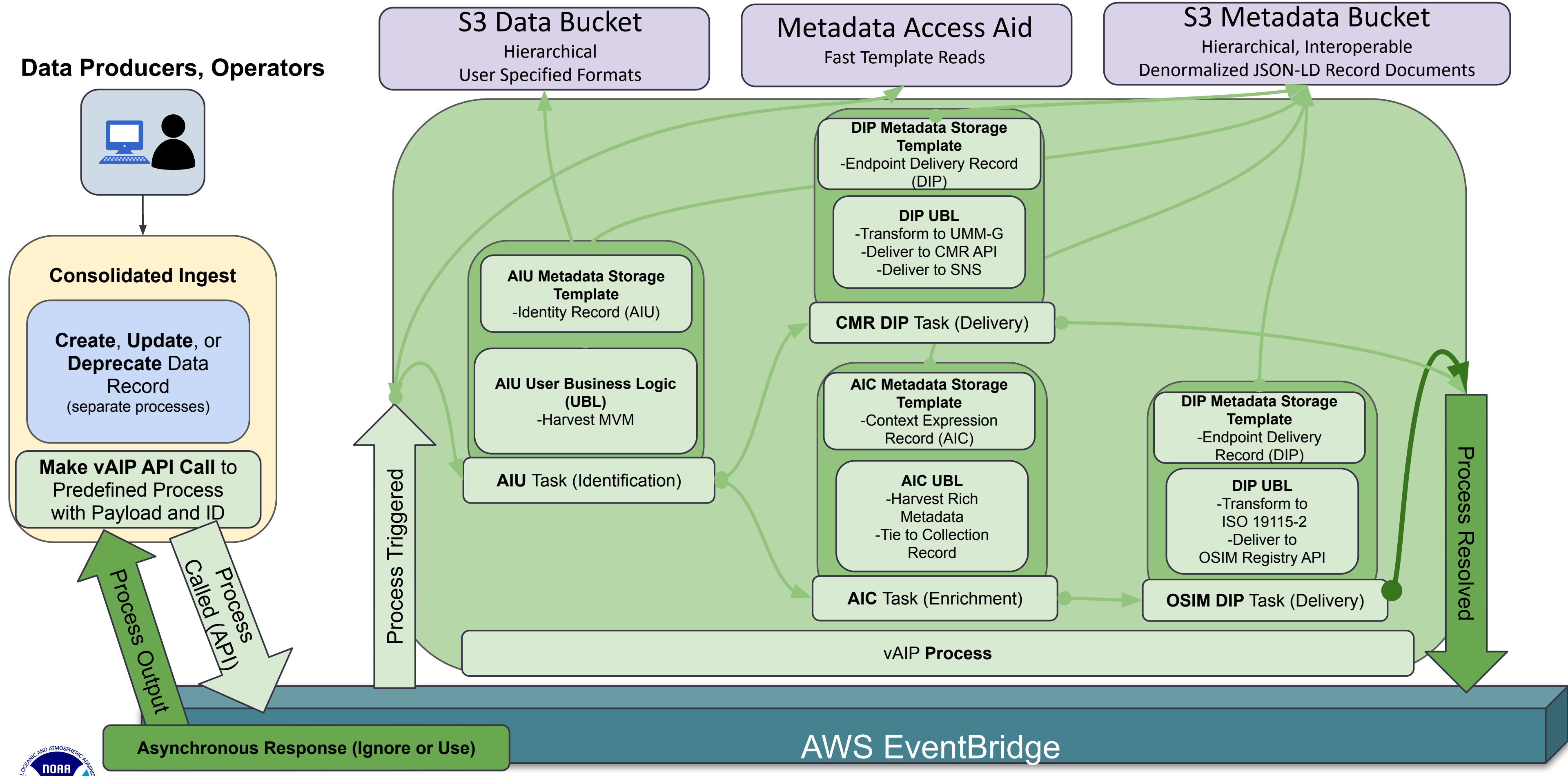


Data Producers, Operators
S3 Data Bucket
Hierarchical
User Specified Formats
Metadata Access Aid
Fast Template Reads
S3 Metadata Bucket
Hierarchical, Interoperable
Denormalized JSON-LD Record Documents
Automated Metadata Record Generation and Storage
Consolidated Ingest
Create, Update, or Deprecate Data Record
(separate processes)
Make vAIP API Call to Predefined Process with Payload and ID
DIP Metadata Storage Template
-Endpoint Delivery Record (DIP)
DIP UBL
-Transform to UMM-G
-Deliver to CMR API
-Deliver to SNS
CMR DIP Task (Delivery)
AIU Metadata Storage Template
-Identity Record (AIU)
AIU User Business Logic (UBL)
-Harvest MVM
AIU Task (Identification)
AIC Metadata Storage Template
-Context Expression Record (AIC)
AIC UBL
-Harvest Rich Metadata
-Tie to Collection Record
AIC Task (Enrichment)
DIP Metadata Storage Template
-Endpoint Delivery Record (DIP)
DIP UBL
-Transform to ISO 19115-2
-Deliver to OSIM Registry API
OSIM DIP Task (Delivery)
Process Triggered
Process Resolved
Process Called (API)
Process Output
vAIP Process
AWS EventBridge



Data Producers, Operators
S3 Data Bucket
Hierarchical
User Specified Formats
Metadata Access Aid
Fast Template Reads
S3 Metadata Bucket
Hierarchical, Interoperable
Denormalized JSON-LD Record Documents
Consolidated Ingest
Create, Update, or Deprecate Data Record
(separate processes)
Make vAIP API Call to Predefined Process with Payload and ID
Process Output
Process Called (API)
Process Triggered
AIU Metadata Storage Template
-Identity Record (AIU)
AIU User Business Logic (UBL)
-Harvest MVM
AIU Task (Identification)
Parallel Task Executions
DIP Metadata Storage Template
-Endpoint Delivery Record (DIP)
DIP UBL
-Transform to UMM-G
-Deliver to CMR API
-Deliver to SNS
CMR DIP Task (Delivery)
AIC Metadata Storage Template
-Context Expression Record (AIC)
AIC UBL
-Harvest Rich Metadata
-Tie to Collection Record
AIC Task (Enrichment)
DIP Metadata Storage Template
-Endpoint Delivery Record (DIP)
DIP UBL
-Transform to ISO 19115-2
-Deliver to OSIM Registry API
OSIM DIP Task (Delivery)
Process Resolved
vAIP Process
AWS EventBridge


S3 Data Bucket
Hierarchical
User Specified Formats
Metadata Access Aid
Fast Template Reads
S3 Metadata Bucket
Hierarchical, Interoperable
Denormalized JSON-LD Record Documents
Data Producers, Operators
Consolidated Ingest
Create, Update, or Deprecate Data Record
(separate processes)
Make vAIP API Call to Predefined Process with Payload and ID
Process Output
Process Called (API)
Process Triggered
AIU Metadata Storage Template
-Identity Record (AIU)
AIU User Business Logic (UBL)
-Harvest MVM
AIU Task (Identification)
DIP Metadata Storage Template
-Endpoint Delivery Record (DIP)
DIP UBL
-Transform to UMM-G
-Deliver to CMR API
-Deliver to SNS
CMR DIP Task (Delivery)
Partial Request Resolution
AIC Metadata Storage Template
-Context Expression Record (AIC)
AIC UBL
-Harvest Rich Metadata
-Tie to Collection Record
AIC Task (Enrichment)
DIP Metadata Storage Template
-Endpoint Delivery Record (DIP)
DIP UBL
-Transform to ISO 19115-2
-Deliver to OSIM Registry API
OSIM DIP Task (Delivery)
Process Resolved
vAIP Process
Asynchronous Response (Ignore or Use)
AWS EventBridge


S3 Data Bucket
Hierarchical
User Specified Formats
Metadata Access Aid
Fast Template Reads
S3 Metadata Bucket
Hierarchical, Interoperable
Denormalized JSON-LD Record Documents
Data Producers, Operators
Consolidated Ingest
Create, Update, or Deprecate Data Record
(separate processes)
Make vAIP API Call to Predefined Process with Payload and ID
Process Output
Process Called (API)
Process Triggered
AIU Metadata Storage Template
-Identity Record (AIU)
AIU User Business Logic (UBL)
-Harvest MVM
AIU Task (Identification)
DIP Metadata Storage Template
-Endpoint Delivery Record (DIP)
DIP UBL
-Transform to UMM-G
-Deliver to CMR API
-Deliver to SNS
CMR DIP Task (Delivery)
AIC Metadata Storage Template
-Context Expression Record (AIC)
AIC UBL
-Harvest Rich Metadata
-Tie to Collection Record
AIC Task (Enrichment)
DIP Metadata Storage Template
-Endpoint Delivery Record (DIP)
DIP UBL
-Transform to ISO 19115-2
-Deliver to OSIM Registry API
OSIM DIP Task (Delivery)
Process Resolved
vAIP Process
AWS EventBridge
Asynchronous Response (Ignore or Use)

Common Service - Ordering System

Access Order Broker - Superset of Order Fulfillment Systems