NARA, the OAIS-RM, and NOAA's Next-Generation Cloud Archive
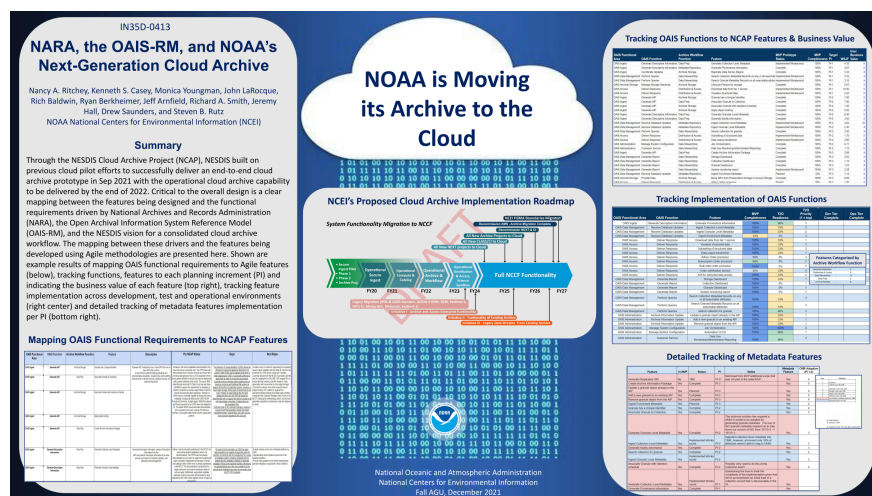
Nancy Ritchey[1], Kenneth Casey[1], Monica Youngman[1], John LaRocque[1], Rich Baldwin[1], Ryan Berkheimer[1], Jeff Arnfield[1], Richard A. Smith[1], Jeremy Hall[1], Drew Saunders[1], and Steven Rutz[1]

[1]NOAA/NCEI

November 24, 2022

Abstract

Through the NESDIS Cloud Archive Project (NCAP), NESDIS is building on previous cloud pilot efforts and making significant progress on delivering an end-to-end cloud archive prototype in late 2021 and operational cloud archive capability in 2022. Critical to the overall design is a clear mapping between the features being designed and the functional requirements driven by National Archives and Records Administration (NARA), the Open Archival Information System Reference Model (OAIS-RM), and the NESDIS vision for a consolidated cloud archive workflow. The mapping between these drivers and the features being developed using Agile methodologies will be reviewed, and a technical summary of NESDIS' plans to implement data archiving using commercial cloud services will be provided.

NARA, the OAIS-RM, and NOAA's Next-Generation Cloud Archive

Nancy A. Ritchey, Kenneth S. Casey, Monica Youngman, John LaRocque, Rich Baldwin, Ryan Berkheimer, Jeff Arnfield, Richard A. Smith, Jeremy Hall, Drew Saunders, and Steven B. Rutz
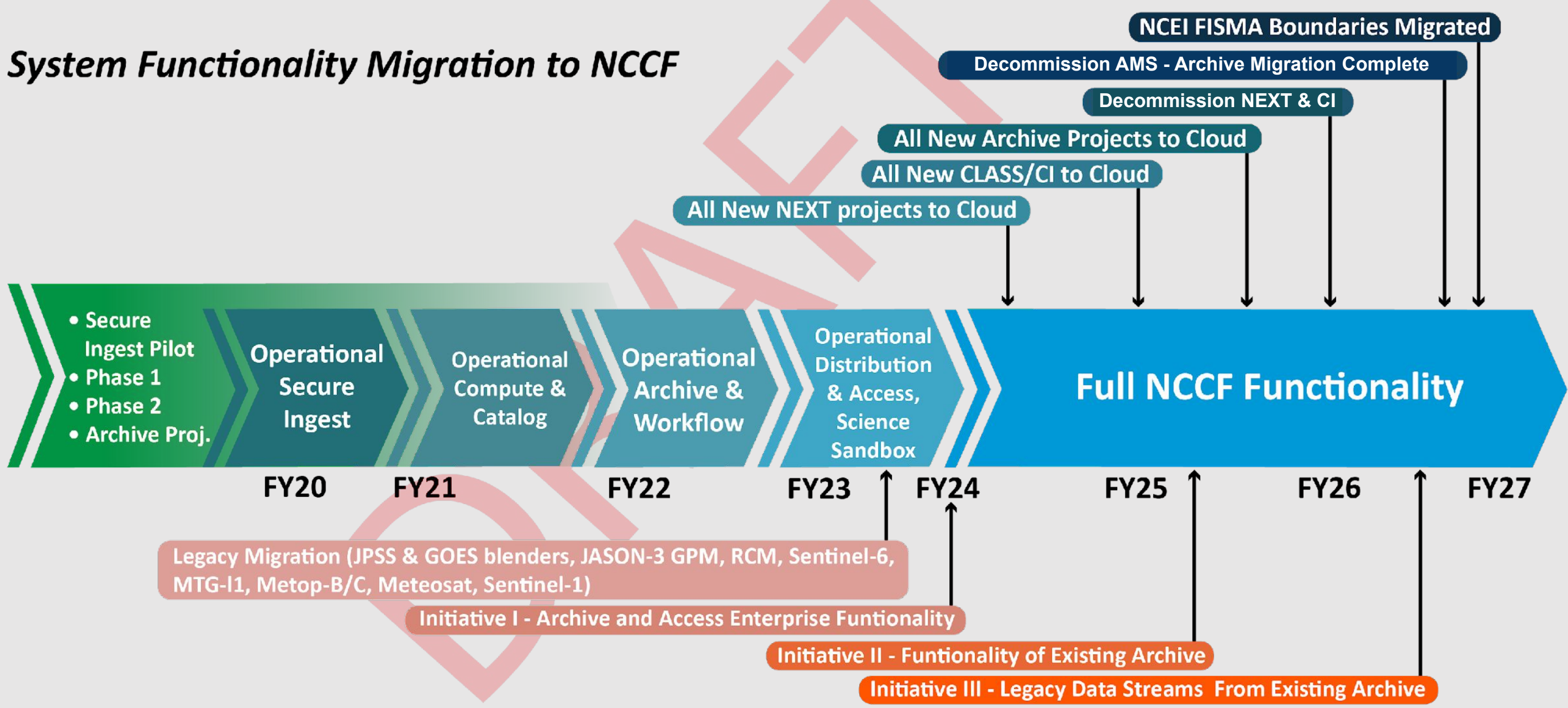NOAA National Centers for Environmental Information (NCEI)

Summary

Through the NESDIS Cloud Archive Project (NCAP), NESDIS built on previous cloud pilot efforts to successfully deliver an end-to-end cloud archive prototype in Sep 2021 with the operational cloud archive capability to be delivered by the end of 2022. Critical to the overall design is a clear mapping between the features being designed and the functional requirements driven by National Archives and Records Administration (NARA), the Open Archival Information System Reference Model (OAIS-RM), and the NESDIS vision for a consolidated cloud archive workflow. The mapping between these drivers and the features being developed using Agile methodologies are presented here. Shown are example results of mapping OAIS functional requirements to Agile features (below), tracking functions, features to each planning increment (PI) and indicating the business value of each feature (top right), tracking feature implementation across development, test and operational environments (right center) and detailed tracking of metadata features implementation per PI (bottom right).

NOAA is Moving its Archive to the Cloud

Mapping OAIS Functional Requirements to NCAP Features

OAIS Functional Area	OAIS Function	Archive Workflow Function	Feature	Description	Pre NCAP Status	Gaps	Next Steps
OAIS Ingest	Generate AIP	Archival Storage	Granule has a Unique Identifier				
OAIS Ingest	Generate AIP	Data Prep	Associate Granule to Collection				
OAIS Ingest	Generate AIP	Archival Storage	Associate Granule with retention schedule				
OAIS Ingest	Generate AIP	Archival Storage	Apply object locking				
OAIS Ingest	Generate AIP	Data Prep	Create Archive Information Package				
OAIS Ingest	Generate Descriptive Information	Data Prep	Generate Collection Level Metadata				
OAIS Ingest	Generate Descriptive Information	Data Prep	Generate Granular Level Metadata				

Tracking OAIS Functions to NCAP Features & Business Value

OAIS Functional Area	OAIS Function	Archive Workflow Function	Feature	MVP Prototype Status	MVP Completeness	Target PI	WSJF	User Business Value
OAIS Ingest	Generate Descriptive Information	Data Prep	Generate Collection Level Metadata	Implemented-Workaround	100%	PI-1	4.33	5
OAIS Ingest	Generate Descriptive Information	Metadata Repository	Generate Provenance information	Complete	100%	PI-1	3.67	5
OAIS Ingest	Coordinate Updates	Archival Storage	Replicate Data Across Region	Complete	100%	PI-1	3.33	4
OAIS Data Management	Perform Queries	Data Stewardship	Search Collection Metadata Records on any or all searchab	Implemented-Workaround	100%	PI-1	3.33	4
OAIS Data Management	Perform Queries	Data Stewardship	Search Granule Metadata Records on all searchable attribu	Implemented-Workaround	100%	PI-1	3.33	4
OAIS Archival Storage	Manage Storage Hierarchy	Archival Storage	Lifecycle Policies for storage	Complete	100%	PI-1	2.67	3
OAIS Access	Deliver Response	Distribution & Access	Download data from tier 1 service	Implemented-Workaround	100%	PI-1	15.00	5
OAIS Access	Deliver Response	Distribution & Access	Visualize structured data	Implemented-Workaround	100%	PI-1	2.20	5
OAIS Ingest	Generate AIP	Archival Storage	Granule has a Unique Identifier	Complete	100%	PI-2	7.00	5
OAIS Ingest	Generate AIP	Data Prep	Associate Granule to Collection	Complete	100%	PI-2	7.00	5
OAIS Ingest	Generate AIP	Archival Storage	Associate Granule with retention schedule	Complete	100%	PI-2	1.50	5
OAIS Ingest	Generate AIP	Archival Storage	Apply object locking	Complete	100%	PI-2	5.50	4
OAIS Ingest	Generate Descriptive Information	Data Prep	Generate Granular Level Metadata	Complete	100%	PI-2	6.50	5
OAIS Ingest	Generate Descriptive Information	Data Prep	Generate fixity information	Complete	100%	PI-2	3.50	2
OAIS Data Management	Receive Database Updates	Metadata Repository	Ingest Collection Level Metadata	Implemented-Workaround	100%	PI-2	4.00	10
OAIS Data Management	Receive Database Updates	Metadata Repository	Ingest Granular Level Metadata	Complete	100%	PI-2	2.40	4
OAIS Data Management	Perform Queries	Data Stewardship	Search collection for granule	Complete	100%	PI-2	3.00	4
OAIS Access	Deliver Response	Distribution & Access	Subsetting of structured data	Implemented-Workaround	100%	PI-2	1.75	5
OAIS Access	Deliver Response	Distribution & Access	Data output transformer	Implemented-Workaround	100%	PI-2	2.00	5
OAIS Administration	Manage System Configuration	Data Stewardship	Job Orchestration	Complete	100%	PI-2	0.71	5
OAIS Administration	Customer Service	Data Stewardship	Data Use Monitoring/Administration/Reporting	Complete	100%	PI-2	1.15	5
OAIS Ingest	Generate AIP	Data Prep	Create Archive Information Package	Complete	100%	PI-3	2.88	4
OAIS Data Management	Generate Report	Data Stewardship	Storage Dashboard	Complete	100%	PI-3	2.50	4
OAIS Data Management	Generate Report	Data Stewardship	Collection Dashboard	Complete	100%	PI-3	1.13	4
OAIS Data Management	Generate Report	Data Stewardship	Granule Dashboard	Complete	100%	PI-3	1.23	4
OAIS Data Management	Generate Report	Data Stewardship	System monitoring report	Implemented-Workaround	100%	PI-3	2.38	4
OAIS Data Management	Receive Database Updates	Metadata Repository	Ingest Enrichment Metadata	Planned	33%	PI-3	1.13	4
OAIS Archival Storage	Provide Data	Archival Storage	Move AIPs from Preservation Storage to Access Storage	Complete	100%	PI-3	0.69	3
OAIS Access	Deliver Response	Distribution & Access	Adhoc Order processor	Started	90%	PI-3	1.00	

NCEI's Proposed Cloud Archive Implementation Roadmap

System Functionality Migration to NCCF

DRAFT

- NCEI FISMA Boundaries Migrated
- Decommission AMS - Archive Migration Complete
- Decommission NEXT & CI
- All New Archive Projects to Cloud
- All New CLASS/CI to Cloud
- All New NEXT projects to Cloud

- Secure Ingest Pilot
- Phase 1
- Phase 2
- Archive Proj.

| Operational Secure Ingest | Operational Compute & Catalog | Operational Archive & Workflow | Operational Distribution & Access, Science Sandbox | Full NCCF Functionality |

FY20 — FY21 — FY22 — FY23 — FY24 — FY25 — FY26 — FY27

Legacy Migration (JPSS & GOES blenders, JASON-3 GPM, RCM, Sentinel-6, MTG-I1, Metop-B/C, Meteosat, Sentinel-1)
Initiative I - Archive and Access Enterprise Funtionality
Initiative II - Funtionality of Existing Archive
Initiative III - Legacy Data Streams From Existing Archive

Tracking Implementation of OAIS Functions

OAIS Functional Area	OAIS Function	Feature	MVP Completeness	T2O Readiness	T2O Priority (1 = top)	Dev Tier Complete	Ops Tier Complete
OAIS Ingest	Generate Descriptive Information	Generate Provenance information	100%	66%	3		
OAIS Data Management	Receive Database Updates	Ingest Collection Level Metadata	100%	33%	2		
OAIS Data Management	Receive Database Updates	Ingest Granular Level Metadata	100%	33%	2		
OAIS Data Management	Receive Database Updates	Ingest Enrichment Metadata	33%	0%	3		
OAIS Access	Deliver Response	Download data from tier 1 service	100%	33%	1		
OAIS Access	Deliver Response	Visualize structured data	100%	33%	3		
OAIS Access	Deliver Response	Subsetting of structured data	100%	33%	3		
OAIS Access	Deliver Response	Data output transformer	100%	33%	3		
OAIS Access	Deliver Response	Adhoc Order processor	90%	0%	1		
OAIS Access	Deliver Response	Subscription Order processor	66%	0%	2		
OAIS Access	Deliver Response	Bulk order order processor	100%	0%	3		
OAIS Access	Deliver Response	Order certification service	90%	33%	3		
OAIS Access	Deliver Response	A/A for restricted data access	100%	33%	2		
OAIS Data Management	Generate Report	Storage Dashboard	100%	33%	1		
OAIS Data Management	Generate Report	Collection Dashboard	100%	0%	1		
OAIS Data Management	Generate Report	Granule Dashboard	100%	0%	1		
OAIS Data Management	Generate Report	System monitoring report	100%				
OAIS Data Management	Perform Queries	Search Collection Metadata Records on any or all searchable attributes	100%	33%	3		
OAIS Data Management	Perform Queries	Search Granule Metadata Records on all searchable attributes	100%	33%	3		
OAIS Data Management	Perform Queries	Search collection for granule	100%	66%	3		
OAIS Administration	Archival Information Update	Update a granule object already in the AIP	100%	33%	1		
OAIS Administration	Archival Information Update	Add a new granule to an existing AIP	100%	33%	1		
OAIS Administration	Archival Information Update	Remove granule object from the AIP	100%	33%	1		
OAIS Administration	Manage System Configuration	Job Orchestration	100%	100%	3		
OAIS Administration	Manage System Configuration	Automation CI/CD	100%	66%	1		
OAIS Administration	Customer Service	Data Use Monitoring/Administration/Reporting	100%	66%	3		

Features Categorized by Archive Workflow Function

Metadata Repository	4
Distribution & Access	9
Data Stewardship	13
Data Prep	5
Archival Storage	6

Detailed Tracking of Metadata Features

Feature	In MVP	Status	PI	Notes	Metadata Feature	CMR Adoption (PI 1-3)
Generate Replicated URI	No	TBD	PI-5*	Removed from MVP additional scope that was not part of the initial MVP.	Yes	0
Create Archive Information Package	Yes	Complete	PI-3		Yes	0
Update a granule object already in the AIP	Yes	Planned	PI-3		Yes	0
Add a new granule to an existing AIP	Yes	Complete	PI-3		Yes	0
Remove granule object from the AIP	Yes	Complete	PI-3		Yes	0
Ingest Enrichment Metadata	Yes	Planned	PI-3		Yes	2
Granule has a Unique Identifier	Yes	Complete	PI-2		Yes	4
Associate Granule to Collection	Yes	Complete	PI-2		Yes	4
Generate Granular Level Metadata	Yes	Complete	PI-2	The technical solution has required a UMM-G model to be adopted for generating granule metadata. The use of ISO granule metadata required us to step down our version of ISO from 19115-2 --> 19115-1.	Yes	3
Ingest Collection Level Metadata	Yes	Implemented-Workaround	PI-2	Ingested collection level metadata into CMR, however, at present only 16% of attributes weren't able to map to UMM.	Yes	3
Generate fixity information	Yes	Complete	PI-2		Yes	0
Search collection for granule	Yes	Complete	PI-2		Yes	0
Ingest Granular Level Metadata	Yes	Implemented-Workaround	PI-2		Yes	3
Associate Granule with retention schedule	Yes	Complete	PI-2	Possibly only need to do this at the Collection level?	Yes	0
Generate Collection Level Metadata	Yes	Implemented-Workaround	PI-1	Questioning the how to track the complexity of the implementation given that we've accomplished an initial load of a collection record that is discoverable in the UI.	Yes	3
Generate Provenance information	Yes	Complete	PI-1		Yes	