



Blu

loud

Piloting innovative services for Marine Research & the Blue Economy

Global Record of Stocks and Fisheries

GRSF

Towards Refreshing GRSF Contents

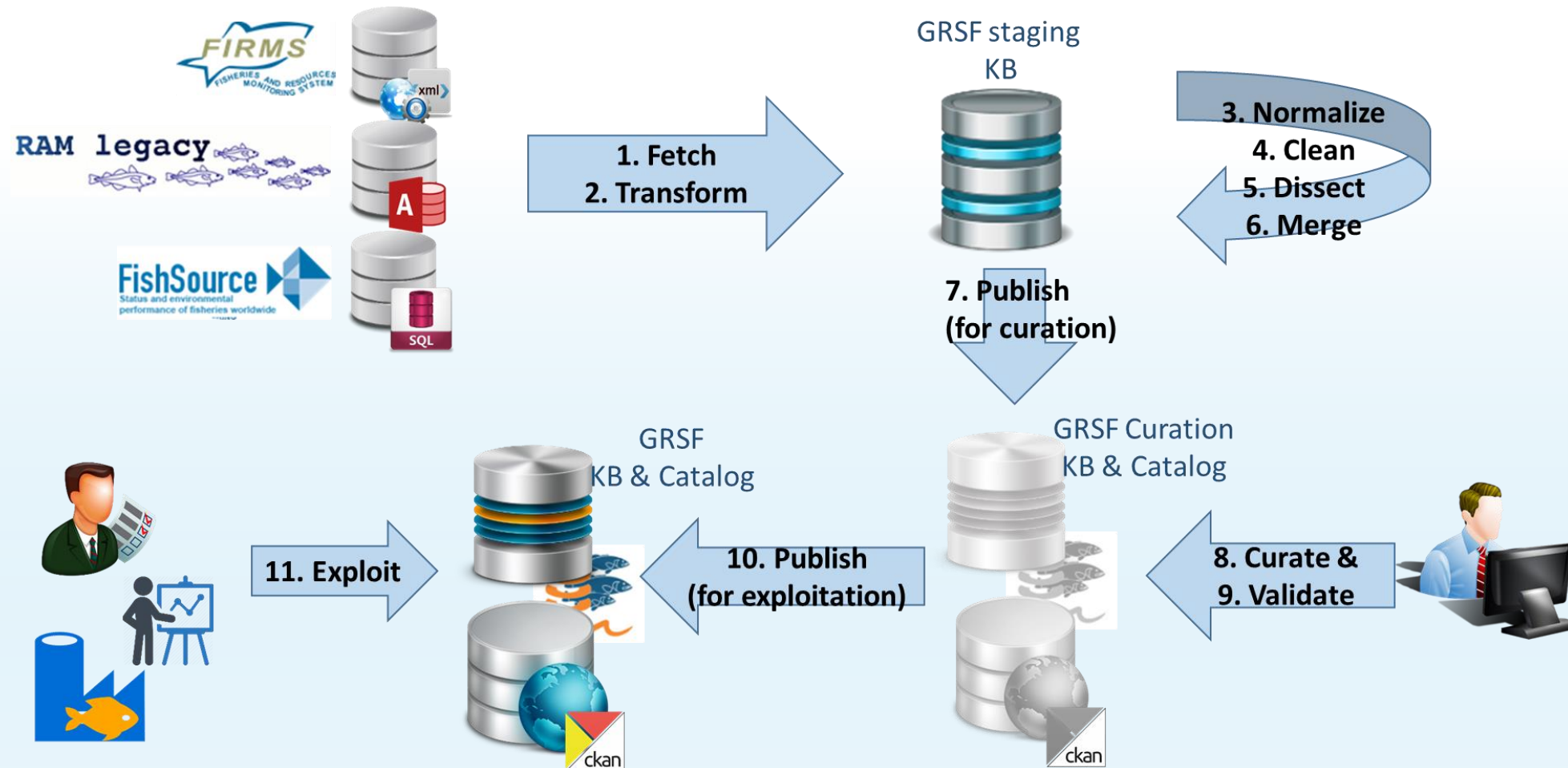


Blue-Cloud has received funding from the European Union's Horizon Programme call BG-07-2019-2020, topic: [A] 2019 - Blue Cloud services, grant Agreement number 862409.

Outline

- GRSF Construction Process
- Sources of Information
- Information to be preserved
- Update Workflow (high-level view)
 - Update Legacy Records
 - Update GRSF Stock Records
 - Update GRSF Fishery Records
- Analytics

GRSF Construction Process



Sources of Information

  	GRSF KB 	GRSF Catalogs 
URL http://firms.fao.org/fishery/xml/resource/10157/en ID WPOLLNAVAR URL https://www.fishsource.org/stock_page/1531	URI & UUID grsf:921bfa9e-d9ba-402b-b11e-fd94a352c5d1	URL & UUID (catalog) http://data.d4science.org/ctlg/GRSF...5f8e0d94-0455-44d5-bde2-92dd8ab8fb20
	← refers to firms.URL ram.ID fishsource.URL catalog.URL catalog.ID → refers to	

Information to be preserved

Legacy Records

-  Catalog URL & UUID (for ADMIN Catalog)

GRSF Records

-  Catalog URL & UUID (for ADMIN and Public Catalogs)
-  Record status (i.e. pending/ approved/ archived/ rejected/ hidden)
-  User Annotations (i.e. message, user, date and time)
-  Merging of records (e.g. a GRSF record constructed from two or more legacy stock records)



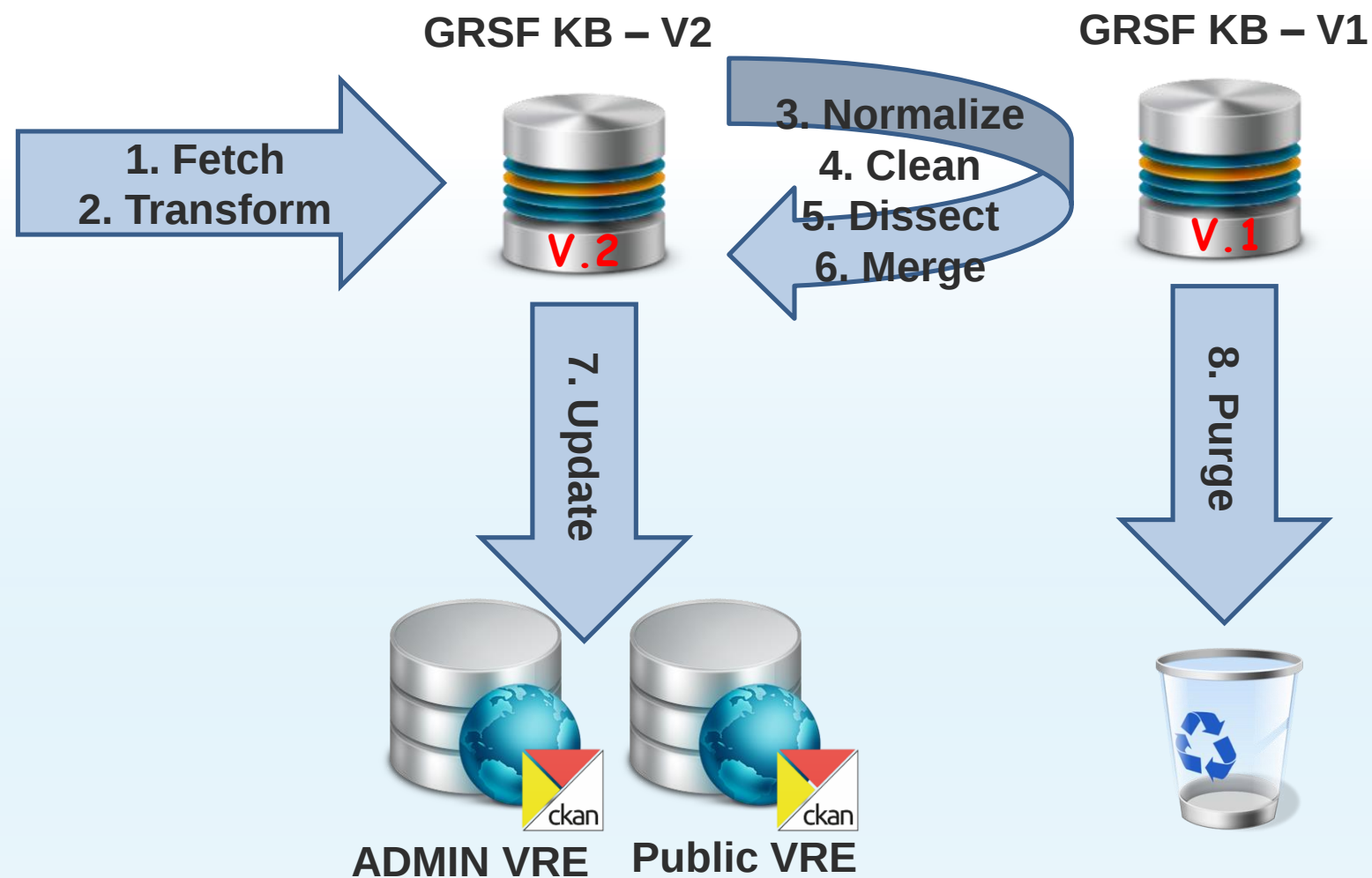
Blu loud



RAM legacy



GRSF Update Workflow



Update Legacy Records

```
PI[]: [{ "source_url" : "http://firms.fao.org/fishery/xml/resource/13548/en",  
        "uuid" : "e9b48cd0-d990-38ec-8dfe-0cd7d1248b0c",  
        "catalogue_url" : "http://data.d4science.org/ctlg/GRSF_Admin/e9b48cd0-d990-38ec-8dfe-0cd7d1248b0c",  
        "catalogue_id" : "7fc81c7a-60fb-40d6-8346-8879a1088c96"  
      }]
```

- PI[] : Export information from GRSF KB V.1
- for each instance in PI[]
 - if ($\exists$ record in GRSF KB V.2 $\rightarrow$ record.source_url == instance.source_url)
 - Update the corresponding record in GRSF KB V.2
- for each record in GRSF KB V.2 (that do not exist in PI[])
 - Generate a UUID
- Update all records in GRSF Admin VRE using GRSF KB V.2

Update Legacy Records – cont'd

- Obsolete Records (published Catalog but they do not exist anymore)
 - Should we remove them?

	Updated	New	Obsolete
FIRMS	929	209	4
RAM	1290	84	1
FishSource	4096	1132	115

Update GRSF Stock Records

```
[{"type" : "stock",  
  "uuid" : "0dcd3d3d-bfa0-32b1-9372-88082d721b54",  
  "catalogue_url" : "http://data.d4science.org/ctlg/GRSF_Admin/0dcd3d3d-bfa0-32b1-9372-88082d721b54",  
  "catalogue_id" : "af2cca19-c80d-4dd5-938f-8ecb3c6cfdcd",  
  "status" : "approved",  
  "semantic_id" : "asfis:TOP+fao:48.3",  
  "source_urls" : [  
    "http://firms.fao.org/fishery/xml/resource/10509/en",  
    "https://www.fishsource.org/stock_page/850"  
  ],  
  "public_catalogue_url" : "http://data.d4science.org/ctlg/GRSF/0dcd3d3d-bfa0-32b1-9372-88082d721b54",  
  "public_catalogue_id" : "58121b02-c089-4685-98d5-46345b089c6b",  
  "annotations" : [  
    {"annotation_admin" : "GRSF Administrator",  
     "annotation_message" : "This is an annotation message",  
     "annotation_time" : "2019-12-27 15:58:12 CET"}  
  ]  
}]
```

- Do we need to preserve more information?
 - Yes (Resource type, SDG flag, connected fishery records)

Update GRSF Stock Records – cont'd

- $PI[]$: Export information from GRSF KB V.1
- for each instance in $PI[]$
 - if ($\exists$ records in GRSF KB V.2 $\rightarrow$ record.source_url == instance.source_url)
 - Construct a GRSF Stock record with these source records from GRSF KB V.2
 - Update the GRSF record in GRSF KB V.2 with information from PI
 - else
 - If(record.status==pending)
 - Remove record from the catalog
 - Else
 - Update record.status to archived with a proper annotation for administrators
- for each legacy record
 - if ($\exists$ record in legacy_records $\rightarrow$ record.grsf_record = null)
 - Construct GRSF Stock record
- Update/Publish GRSF Stock records in GRSF ADMIN VRE
- Update GRSF Stock records in GRSF Public VRE

Update GRSF Stock Records – cont'd

- Obsolete GRSF Stock records
 - Should we remove them ?
 - All? / Pending? / Approved?

Type	Records
Refreshed	2,778
New	439
Obsolete	72



Status	Records	FIRM S	RA M	Fish Sour ce
Pending	64		1	63
Approved	8	3		5
Archived	0			
Rejected	0			
Hidden	0			

Update GRSF Fishery Records

```
[{ "type" : "fishery",
  "uuid" : "9af7c61b-81f0-3c4f-aea9-19f77a9568b4",
  "catalogue_url" : "http://data.d4science.org/ctlg/GRSF_Admin/9af7c61b-81f0-3c4f-aea9-19f77a9568b4",
  "catalogue_id" : "50b48c9d-f133-4c4c-8eac-3323a3bb76b5",
  "status" : "approved",
  "semantic_id" : "asfis:COD+fao:27.5.B;fao:27.6.A+authority:INT:EC+iso3:GBR+isscfg:03.12",
  "source_urls" : [ "https://www.fishsource.org/fishery_page/4166" ],
PI[]: "public_catalogue_url" : "http://data.d4science.org/ctlg/GRSF/9af7c61b-81f0-3c4f-aea9-19f77a9568b4",
      "public_catalogue_id" : "d07e026a-d9c1-40c5-9acc-823dd8bcacfa",
      "annotations" : [
        {"annotation_admin" : "GRSF Administrator",
         "annotation_message" : "This is an annotation message",
         "annotation_time" : "2019-12-27 15:58:12 CET"}
      ]
}]
```

- Do we need to preserve more information?
 - Yes (Resource type, SDG flag, connected fishery records)

Update GRSF Fishery Records – cont'd

- $PI[]$: Export information from GRSF KB V.1
- $FR[]$: GRSF Fishery records generated from GRSF KB V.2
- for each record in $FR[]$
 - If $(\exists \text{ instance in } PI[] \rightarrow \text{record.source.source_url} == \text{instance.source_url} \ \&\& \ \text{record.semantic_id.matches}(\text{instance.semantic_id}))$
 - Update record with the corresponding information in GRSF KB V.2
- for each instance in $PI[]$ (that $\text{instance.source_url} == \text{null}$)
 - If($\text{instance.status} == \text{pending}$)
 - Remove instance from the catalog
 - Else
 - Update instance.status to archived with a proper annotation for administrators
- Update/Publish GRSF Stock records in GRSF ADMIN VRE
- Update GRSF Stock records in GRSF Public VRE

Update GRSF Fishery Records – cont'd

- Semantic ID match algorithm
 - <species> + <areasList> + <managementEntitiesList> + <flagState> + <fishingGear>
 - Must be exactly the same (if not empty):
 - <species>, <flagstate>, <fishingGear>
 - Must contain:
 - <managementEntitiesList>, <areasList>
- Examples
 - asfis:COD+fao:27.5.B+authority:INT:EC+iso3:GBR+ ==
asfis:COD+fao:27.5.B+authority:INT:EC+iso3:GBR+**isscfg:03.12**
 - asfis:COD+fao:27.5.B+authority:INT:EC+iso3:GBR+isscfg:03.12 ==
asfis:COD+fao:27.5.B+authority:INT:EC+iso3:GBR+isscfg:03.12
 - asfis:COD+fao:27.5.B+authority:INT:EC+iso3:GBR+isscfg:03.12 ==
asfis:COD+fao:27.5.B+authority:INT:EC;**authority:NAT:GBR**+iso3:GBR+isscfg:03.12

Update GRSF Stock Records – cont'd

- Obsolete GRSF Fishery records
 - Should we remove them ?
 - All? / Pending? / Approved?

Type	Records
Refreshed	3580
New	3731
Obsolete	3288



Status	Records	FIRMS	FishSource
Pending	3182	215	2967
Approved	102	0	102
Archived			
Rejected	4	0	4
Hidden			

Workflow Analytics

Action	Time	Tools/Services
Harvesting & Normalization	1 hour	MatWare
Data Transformation	15 minutes	X3ML Engine
Data Ingestion (GRSF KB)	1 hour	grsf-services-transform
Data Reconciliation (GRSF KB)	30 minutes	grsf-services-transform
Legacy Records Construction	30 minutes	grsf-services-core
GRSF Stock Records Construction	2 hours	grsf-services-core
GRSF Fishery Records Construction	3 hours	grsf-services-core
Update GRSF Admin VRE		grsf-services-core
Update GRSF Public VRE		grsf-services-core



Blu

loud

Piloting innovative services for Marine Research & the Blue Economy

Thank you

