



SeaDataCloud

CDI Replication: The new ingestion process

Peter Thijsse, Bert Broeren, Dick Schaap, Paolo
D'Onorio de Meo, Merret Buurman, and others

The data ingestion (to the cloud)

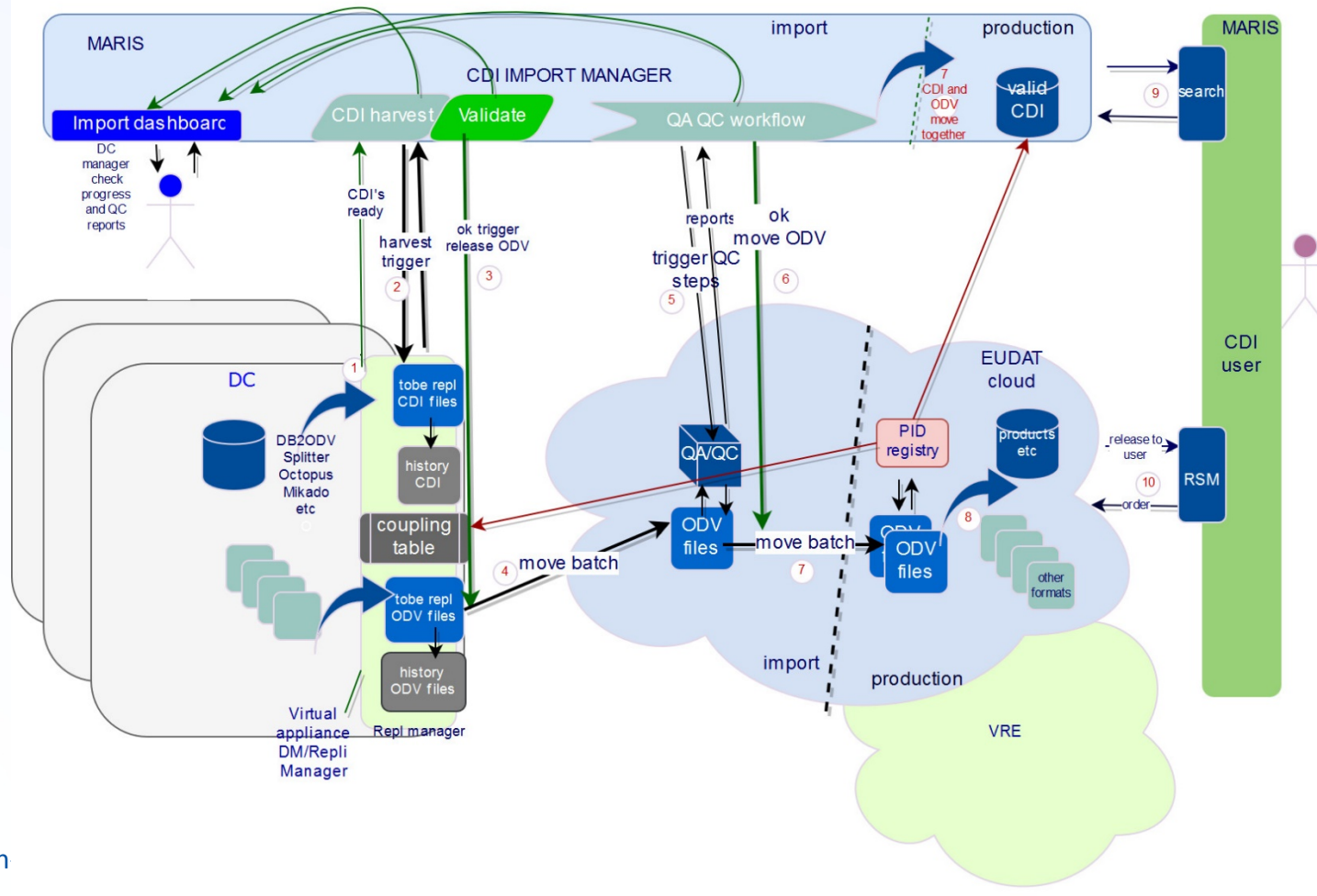
demonstration of process, live and via screens

Main components

- Replication manager
- Import manager
- EUDATs HTTP API
- B2SAFE (Cloud storage)

Overview **unrestricted** dataflow

CDI replication: Unrestricted data
- Very simplified! -



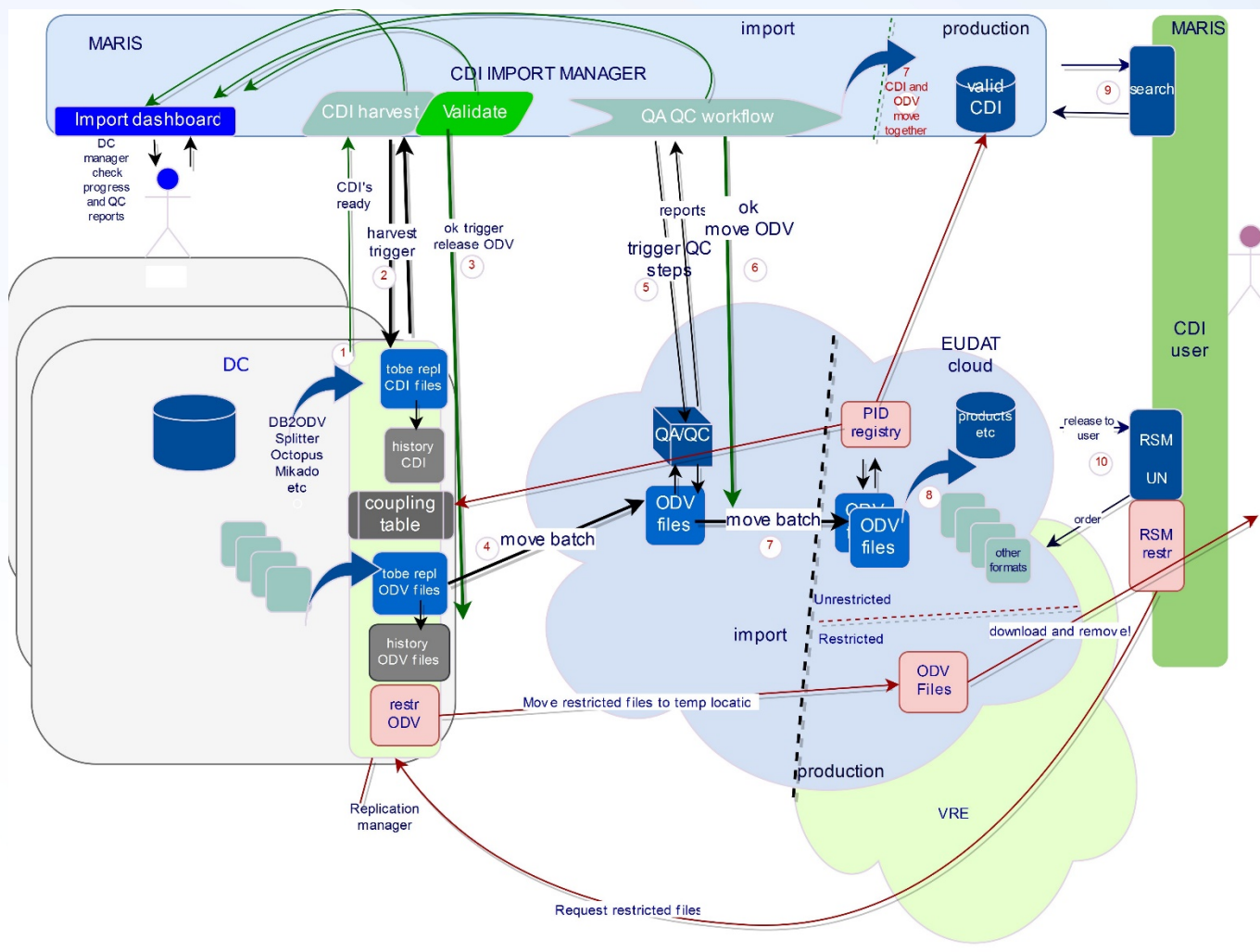
Remarks unrestricted

The main solution is designed for Unrestricted (UN/LS) data where the process of QA/QC and file format conversions in the cloud are involved.

Important steps:

- The CDI import manager is guiding the process of CDI and ODV release and QA/QC towards production.
- At data center there is production of CDI and ODV (or one other accepted exchange data format) using the stack of tools produced at IFREMER: DB2ODV, Splitter, Mikado, Octopus etc.
- Replication Manager at the Data Center is triggered by import manager to:
 - Move a batch of CDI files to CDI central server
 - Follow up requests to move a batch of ODV files triggered by import manager. One batch at a time handled by the system.
- After moving a batch it is archived by the RM (history of full situation).
- The EUDAT cloud consists of a area for Import and an area of Production. During import QA/QC processes will be triggered by the CDI import manager after which reports flow back. After an ok the datasets will be made available in production. Again moved after a trigger from the import manager.
- When a dataset moves to production the PID is reported back to import manager and DC (coupling table)
- The DC data manager (and masters) can follow progress and if necessary take actions to get to a next stage, or e.g. discard the set and release a new updated batch.

Overview **restricted** dataflow



Remarks restricted

Replication Manager at the Data Center is for restricted data triggered by RSM to:

- Follow up requests for releasing and moving a set of restricted data
 - Only when allowed by the DC via confirmation in the RSM
 - Directly to a “user section” in a restricted part of the cloud where it will be available to the user to download it for limited time.
 - After download it will be removed.

Replication of RS data is a sensitive step:

- Trust is needed from the DC in the secure part of the cloud to temporarily store the data
- But, working with the same “trigger and move” process/protocols for each release of data makes it simpler
- The set will be removed in the cloud once downloaded, so it is ready for the next batch.

6 steps in the unrestricted data ingestion process

1. Authentication
2. Publishing metadata
3. Checking metadata
4. Uploading a databatch
5. Unzip and run data checks
6. Move validated metadata and data in production

Step 1: Authentication

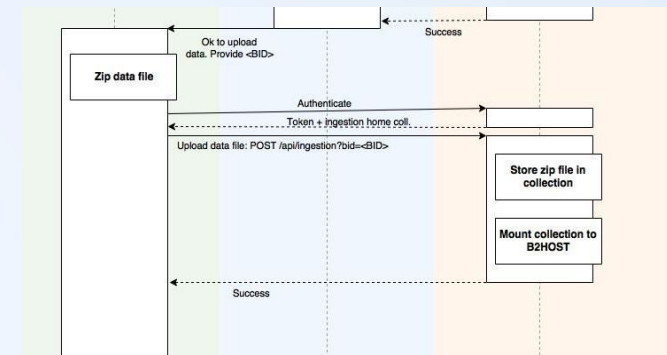
The problem:

RM needs to access the Import Manager and the EUDAT B2STAGE HTTP-API through a software (RM).

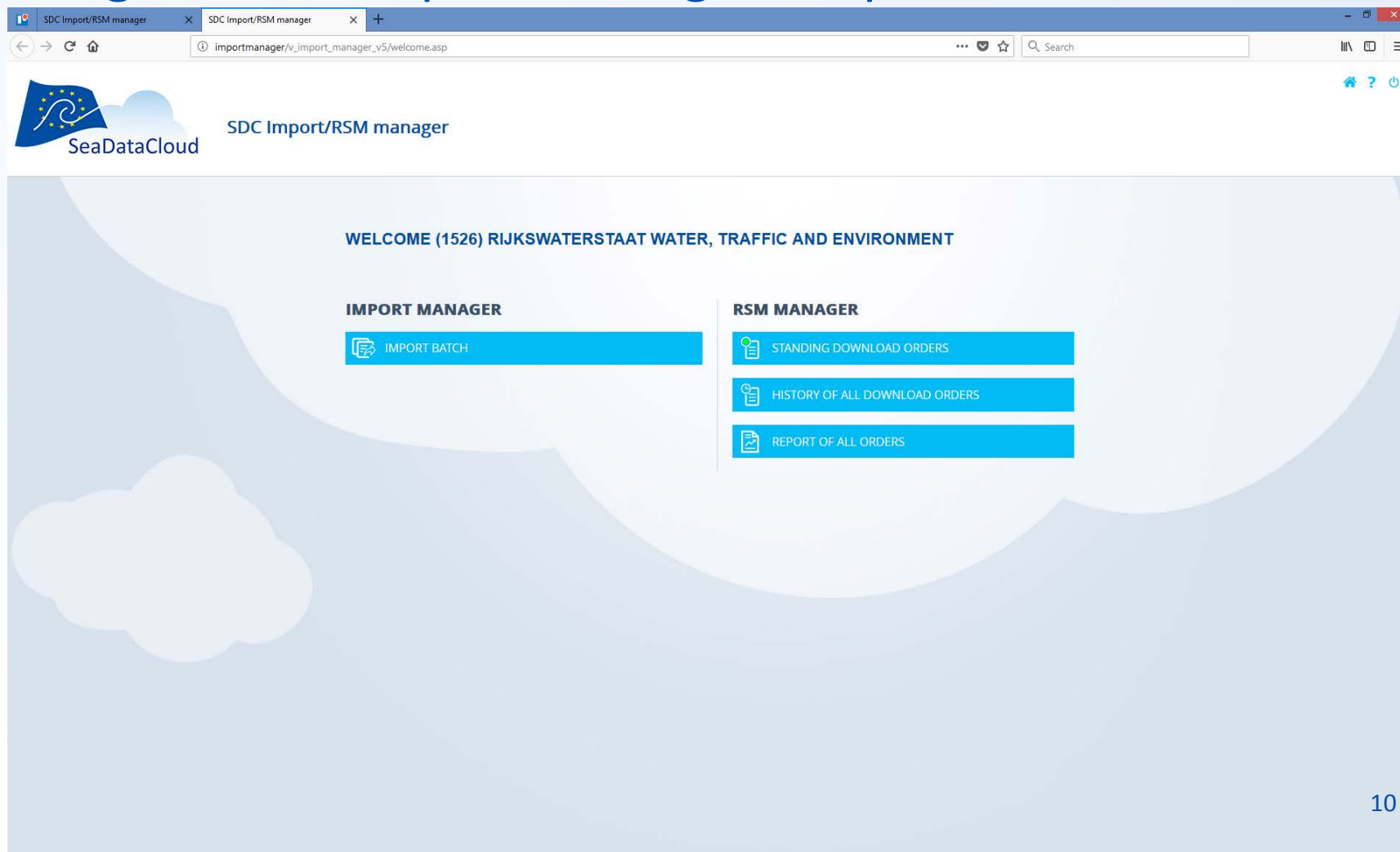
Data center contacts needs to access the IM dashboard.

The solution

1. Every RM registers to the B2ACCESS web site. MARIS prepared neutral “@sdn-taskmanager.org” addresses. These are forwarded to the RSM contacts marineID’s.
2. DC configures the RM to use the credentials
3. RM calls /auth/b2accessproxy with credentials
4. Data center contacts linked to the IM via their MarineID can contact the IM dashboard



Login to the Import Manager for process overview



The screenshot shows a web browser window with two tabs, both labeled "SDC Import/RSM manager". The address bar displays the URL "importmanager/v_import_manager_v5/welcome.asp". The page header includes the SeaDataCloud logo and the text "SDC Import/RSM manager". The main content area features a welcome message: "WELCOME (1526) RIJKSWATERSTAAT WATER, TRAFFIC AND ENVIRONMENT". Below this, there are two sections: "IMPORT MANAGER" and "RSM MANAGER". The "IMPORT MANAGER" section contains a single button labeled "IMPORT BATCH". The "RSM MANAGER" section contains three buttons: "STANDING DOWNLOAD ORDERS", "HISTORY OF ALL DOWNLOAD ORDERS", and "REPORT OF ALL ORDERS".

SDC Import/RSM manager

importmanager/v_import_manager_v5/welcome.asp

SeaDataCloud SDC Import/RSM manager

WELCOME (1526) RIJKSWATERSTAAT WATER, TRAFFIC AND ENVIRONMENT

IMPORT MANAGER

- IMPORT BATCH

RSM MANAGER

- STANDING DOWNLOAD ORDERS
- HISTORY OF ALL DOWNLOAD ORDERS
- REPORT OF ALL ORDERS

Step 2: Publishing the CDI metadata

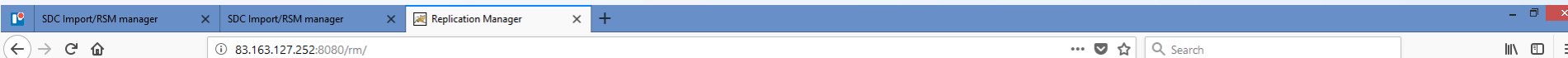
The problem

- RM needs to notify the Import Manager a new batch is ready (RM)
- IM will download when ready

The solution

1. DC manager puts one or more batches in publish-directory
2. RM triggers
3. IM downloads the CDI metadata batch
4. One batch at a time in the process

View of RM dashboard



Replication Manager Dashboard

[Summary](#)[Batches History](#)[LOCAL_CDI_IDS](#)

2018/05/23 16:52

Maintenance

[Logs](#)[External resources](#)[UPDATE](#)[Coupling Table Checks](#)[LAUNCH CHECKER](#)[Database](#)[BACKUP](#)

Configuration

RM configuration is valid

Automatic parameters

parameter	value	status
objectdb_path	C:\Program Files\Apache Software Foundation\Tomcat 8.5\webapps\rm\db\batchentity.odt	OK
harvestPath_path	C:\Program Files\Apache Software Foundation\Tomcat 8.5\webapps\rm\HARVEST	OK

Custom parameters

parameter	value	status
edmo_code	1526	OK
test_mode	true	OK
proxy_host	null	null
proxy_port	-1	null
dataTranslationFile_path	dataTranslationFile_path	OK

RM: Go to “batches history” to select

creation date			name
<input type="checkbox"/>	2018-06-05 15:05:36		data7.zip
<input type="checkbox"/>	2018-06-14 16:59:38		demo10.zip
<input type="checkbox"/>	2018-06-14 16:59:43		demo11.zip
<input type="checkbox"/>	2018-06-14 16:59:48		demo12.zip
<input type="checkbox"/>	2018-06-14 16:59:52		demo13.zip
<input type="checkbox"/>	2018-06-14 16:59:55		demo14.zip
<input type="checkbox"/>	2018-06-14 17:00:00		demo15.zip
<input checked="" type="checkbox"/>	2018-06-14 16:59:15		demo5.zip
<input type="checkbox"/>	2018-06-14 16:59:09		demo6.zip
<input type="checkbox"/>	2018-06-14 16:59:24		demo7.zip
<input type="checkbox"/>	2018-06-14 16:59:28		demo8.zip
<input type="checkbox"/>	2018-06-14 16:59:33		demo9.zip

Current batch

name	batch_number	request_id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
demo4		939cb7b3-5da5-46c6-95a5-dbed02ef610d	2018-06-19 17:15:00	[INGESTION_PENDING] Ingestion started - in process	[IN_HARVEST_ZONE] Moved CDIs batch to HARVEST directory	[IN_QUEUE_ZONE] Moved to Data queue directory	demo4 (15 files)

Batches in queue

RM dashboard: CDI Batch ready notification submitted

Import Batch		Replication Manager		83.163.127.252:8080/rm/BatchesHistory		Search	
<input type="checkbox"/>	2018-06-14 16:59:48			demo12.zip			
<input type="checkbox"/>	2018-06-14 16:59:52			demo13.zip			
<input type="checkbox"/>	2018-06-14 16:59:55			demo14.zip			
<input type="checkbox"/>	2018-06-14 17:00:00			demo15.zip			
<input type="checkbox"/>	2018-06-14 16:59:09			demo6.zip			
<input type="checkbox"/>	2018-06-14 16:59:24			demo7.zip			
<input type="checkbox"/>	2018-06-14 16:59:28			demo8.zip			
<input type="checkbox"/>	2018-06-14 16:59:33			demo9.zip			

Current batch

name	batch_number	request_id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
demo4		939cb7b3-5da5-46c6-95a5-dbed02ef610d	2018-06-19 17:15:00	[INGESTION_PENDING] Ingestion started - in process	[IN_HARVEST_ZONE] Moved CDIs batch to HARVEST directory	[IN_QUEUE_ZONE] Moved to Data queue directory	demo4 (15 files)

Batches in queue

name	batch_number	request_id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
demo5			2018-06-20 10:35:35	[IN_QUEUE_READY] In queue, CDIs and data ready in queue	[CHECK_READY] Check is done	[IN_QUEUE_ZONE] Moved to Data queue directory	demo5 (15 files)

Batches cancelled

name	batch number	request id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
------	--------------	------------	-----------	---------------------	-------------------	-------------------	------------

IM: Notification of CDI batch received, waiting for download

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

Free search

SEARCH CLEAR

Found 19 Show (1-10)

Key	Input_date	Last_update	Batch	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
325	19/6/2018 17:16 PM	19/6/2018 17:16 PM		(1000) Batch created waiting for metadata download	15	0	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	
310	6/6/2018 12:34 PM	7/6/2018 10:30 AM	True	(2180) Removed from import database	14	14	-	

IM: View into the batch specs

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCHCLEAR

CURRENT ACTIONSTATUS OVERVIEWMETADATA LOGFILES

ResultsFound 19Show 1UpdateClear

Key:325Created:19/6/2018 17:16:07Last update:19/6/2018 17:16:07

Start_api_request_n_code	3814
Batch_status	(1000) Batch created waiting for metadata download
Batch_status	
Batch_test	1
Metadata_path_name	
Metadata_zip_file_name	demo4.zip
Metadata_files_count	15
Metadata_files_ok	0
Metadata_files_error	0
Metadata_duplicates_identifiers	0
Metadata_duplicates_batch	0
Metadata_duplicates_production	0
Metadata_duplicates_batch_log	
Metadata_duplicates_production_log	
Comment	

IM: Batch now received, ready to load

Import Batch
Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment
Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 19 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
325	19/6/2018 17:16 PM	20/6/2018 10:39 AM	True	(1100) Add to loading database	15	15	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	
310	6/6/2018 12:34 PM	7/6/2018 10:30 AM	True	(2180) Removed from import database	14	14	-	

RM: Batch received by IM - Now indicates batchnumber

Import Batch		Replication Manager			
		83.163.127.252:8080/rm/BatchesHistory		Search	
<input type="checkbox"/>	2018-06-14 16:59:48			demo12.zip	
<input type="checkbox"/>	2018-06-14 16:59:52			demo13.zip	
<input type="checkbox"/>	2018-06-14 16:59:55			demo14.zip	
<input type="checkbox"/>	2018-06-14 17:00:00			demo15.zip	
<input type="checkbox"/>	2018-06-14 16:59:09			demo6.zip	
<input type="checkbox"/>	2018-06-14 16:59:24			demo7.zip	
<input type="checkbox"/>	2018-06-14 16:59:28			demo8.zip	
<input type="checkbox"/>	2018-06-14 16:59:33			demo9.zip	

Current batch

name	batch_number	request_id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
demo4	325	939cb7b3-5da5-46c6-95a5-dbed02ef610d	2018-06-19 17:15:00	[INGESTION_PENDING] Ingestion started - in process	[ARCHIVED] CDIs batch harvested and archived	[IN_QUEUE_ZONE] Moved to Data queue directory	demo4 (15 files)

Batches in queue

name	batch_number	request_id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
demo5			2018-06-20 10:35:35	[IN_QUEUE_READY] In queue, CDIs and data ready in queue	[CHECK_READY] Check is done	[IN_QUEUE_ZONE] Moved to Data queue directory	demo5 (15 files)

Batches cancelled

name	batch number	request id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
------	--------------	------------	-----------	---------------------	-------------------	-------------------	------------

Step 3: Checking the CDI metadata

The problem

- Import Manager to check the CDI XML and its content
- RM manager to check also visually

The solution

1. Unzip the batch
2. Run sequence of checks (same as current)
3. All logs added to IM CMS
4. Metadata loaded to import database
5. RM notified visual check in import is needed
6. RM manager decides which files continue to next steps

IM: Batch checked and ready to load in import

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment











Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 19 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
325	19/6/2018 17:16 PM	20/6/2018 10:41 AM	True	(2100) Add to import database	15	15	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	
310	6/6/2018 12:34 PM	7/6/2018 10:30 AM	True	(2180) Removed from import database	14	14	-	

IM: History of steps in process passed

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH

CLEAR

CURRENT ACTION

STATUS OVERVIEW

METADATA LOGFILES

Results

Found 19

Show 1

Update

Clear

Key:325 Created:19/6/2018 17:16:07 Last update:20/6/2018 10:41:06

Batch_status_history

batch_status 2100; 2018-06-20T10:41:05.730
batch_status 1120; 2018-06-20T10:41:05.713
batch_status 1110; 2018-06-20T10:41:00.120
batch_status 1100; 2018-06-20T10:39:47.160
batch_status 1020; 2018-06-20T10:39:47.143

API_requests and errors

Date	API function	Errors
20180620T08:06:14	metadata_to_load_database_ready	no errors
20180620T08:06:14	metadata_to_load_database	no errors
20180620T08:06:55	metadata_download_ready	no errors
20180619T17:15:14	metadata_download	no errors

IM: Status progress and logs XML validation available

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

CURRENT ACTION STATUS OVERVIEW METADATA LOGFILES

Results Found 19 Show 1 Update Clear

Key:325 Created:19/6/2018 17:16:07 Last update:20/6/2018 10:41:06

Import log files

```
2018-06-20 10:40:08
=====
-- start of program 6/20/2018 10:40:08
=====
2018-06-20 10:40:08 Batch.number = 325
2018-06-20 10:40:08 Batch.Edmo_code = 1526
2018-06-20 10:40:08 Batch.Directory = e:\internet\seadatanet\CDI_V5_bestanden_aangeleverd\1526\1526_325
2018-06-20 10:40:08 Batch.files = 0
2018-06-20 10:40:08 Batch.delete = False
2018-06-20 10:40:08
2018-06-20 10:40:13 =====
2018-06-20 10:40:13 Number of records for duplicate check = 15
2018-06-20 10:40:13 =====
2018-06-20 10:40:14
2018-06-20 10:40:14 -----
2018-06-20 10:40:14
2018-06-20 10:40:14 Files Inserted = '15'
2018-06-20 10:40:14 Files error = '0'
2018-06-20 10:40:14 Batch duplicates (identifiers) = '0'
2018-06-20 10:40:14 Potential Batch duplicates = '0'
2018-06-20 10:40:14 Potential Batch Production duplicates = '0'
2018-06-20 10:40:14
```

IM: Move to import database and interface (programme takes a while)

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment










Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 19 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
325	19/6/2018 17:16 PM	20/6/2018 10:43 AM	True	(2120) Addition to import database completed	15	15	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

IM: All ok, ready to load to webinterface (build up tables, elastic search, etc)

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 19 Show (1-10)

Key -	Input_date -	Last_update -	Batch_test -	Batch_status -	Metadata_files_count -	Metadata_files_ok -	N_code -	Edit
325	19/6/2018 17:16 PM	20/6/2018 10:43 AM	True	(2200) Add to import web interface	15	15	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	
310	6/6/2018 12:34 PM	7/6/2018 10:30 AM	True	(2180) Removed from import database	14	14	-	

IM: Almost ready ... (preparing links)

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 19 Show (1-10)

Key -	Input_date -	Last_update -	Batch_test -	Batch_status -	Metadata_files_count -	Metadata_files_ok -	N_code -	Edit
325	19/6/2018 17:16 PM	20/6/2018 10:47 AM	True	(2220) Addition to import web interface completed	15	15	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

IM: Ready to visually check logs and interface (only human action)

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment










Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 19 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
325	19/6/2018 17:16 PM	20/6/2018 10:51 AM	True	(2300) Wait for human visual check in import database Import_database	15	15	link	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

RM: Notification calls datamanager to action

Import Batch

Replication Manager

83.163.127.252:8080/rm/BatchesHistory

Search

<input type="checkbox"/>	2018-06-14 16:59:48	demo12.zip
<input type="checkbox"/>	2018-06-14 16:59:52	demo13.zip
<input type="checkbox"/>	2018-06-14 16:59:55	demo14.zip
<input type="checkbox"/>	2018-06-14 17:00:00	demo15.zip
<input type="checkbox"/>	2018-06-14 16:59:09	demo6.zip
<input type="checkbox"/>	2018-06-14 16:59:24	demo7.zip
<input type="checkbox"/>	2018-06-14 16:59:28	demo8.zip
<input type="checkbox"/>	2018-06-14 16:59:33	demo9.zip

Current batch

name	batch_number	request_id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
demo4	325	939cb7b3-5da5-46c8-95e5-dbed02ef610d	2018-06-14 16:59:33	CMS [HUMAN_ACTIVITY_METADATA] Human action concerning metadata is required on IM CMS	[ARCHIVED] CDIs batch harvested and archived	[IN_QUEUE_ZONE] Moved to Data queue directory	demo4 (15 files)

Batches in queue

name	batch_number	request_id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
demo5			2018-06-20 10:35:35	[IN_QUEUE_READY] In queue, CDIs and data ready in queue	[CHECK_READY] Check is done	[IN_QUEUE_ZONE] Moved to Data queue directory	demo5 (15 files)

Batches cancelled

name	batch_number	request_id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
------	--------------	------------	-----------	---------------------	-------------------	-------------------	------------

IM: Visual check - Individual records can be removed (check live!)


Import Batch

Seadatanet V5

Replication Manager

cdi_import_test.maris.nl/v_cdi_v5/welcome.php?query=325&query_code={F581FF35-B152-405D-8426-420428B2FF70}

Search



PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

PETER THIJSSE

DELETE RECORDS 0

NEW SEARCH

SEARCH RESULTS

SUMMARY

TIMESERIES

Filter search

You searched for:

Batch number:

325

Reset all

INPUT FIELDS

Free search ?

Date search ?

From

yyyymmdd

To

yyyymmdd

Geographic search ?

North

West

East

South

SEARCH

RESET

SEA REGIONS

North Sea (15)

Northeast Atlantic Ocean (40W) (15)

Western Scheldt (14)

DISCIPLINE (P08)

Physical oceanography (15)

PARAMETER GROUP (P03)

Water column temperature a... (15)

DISCOVERY PARAMETER (P02)

20

100

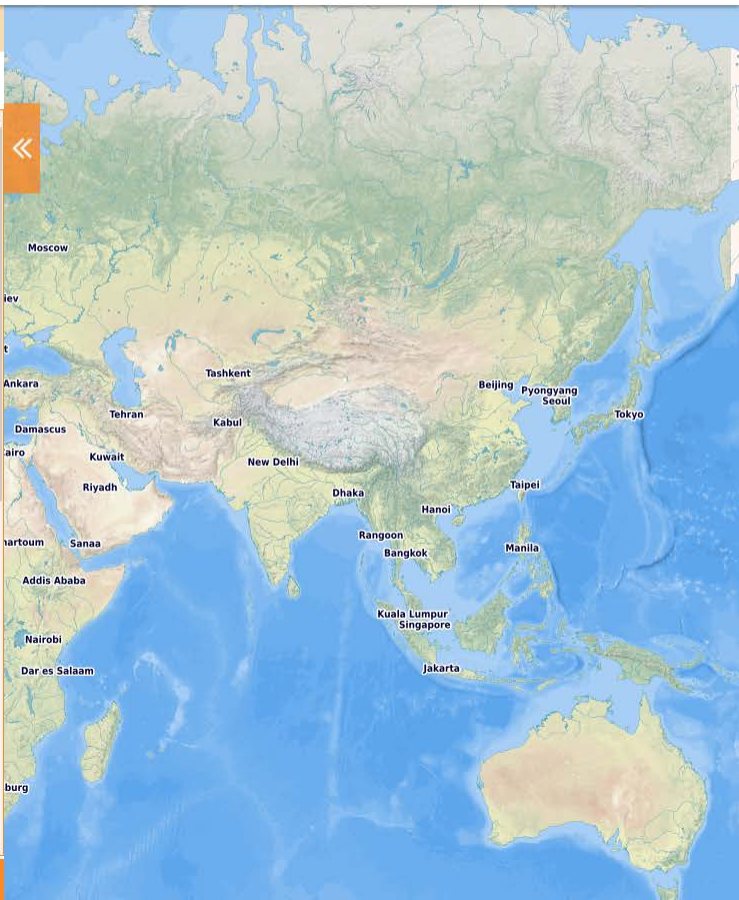
1000

found 15 | show (1 - 15) | Prev | Next

	Dataset name	Country	Start date	Instrument / gear type
<input type="checkbox"/>	Measurements at Bath in 1913 by RIKZMON_TEMP using THER MMTR	Netherlands	19130101	water temperature sensor
<input type="checkbox"/>	Measurements at Bath in 1919 by RIKZMON_TEMP using THER MMTR	Netherlands	19190101	water temperature sensor
<input type="checkbox"/>	Measurements at Bath in 1925 by RIKZMON_TEMP using THER MMTR	Netherlands	19250101	water temperature sensor
<input type="checkbox"/>	Measurements at Bath in 1931 by RIKZMON_TEMP using THER MMTR	Netherlands	19310101	water temperature sensor
<input type="checkbox"/>	Measurements at Bath in 1937 by RIKZMON_TEMP using THER MMTR	Netherlands	19370101	water temperature sensor
<input type="checkbox"/>	Measurements at Bath in 1943 by RIKZMON_TEMP using THER MMTR	Netherlands	19430101	water temperature sensor
<input type="checkbox"/>	Measurements at Bath in 1949 by RIKZMON_TEMP using THER MMTR	Netherlands	19490101	water temperature sensor
<input type="checkbox"/>	Measurements at Bath in 1955 by RIKZMON_TEMP using THER MMTR	Netherlands	19550101	water temperature sensor

EXPORT RESULT

SAVE QUERY



IM: RM manager indicates check is done (go on or cancel batch)

Import Batch Seadatanet V5 Replication Manager +

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4 110% Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

CURRENT ACTION STATUS OVERVIEW METADATA LOGFILES Results Found 19 Show 1 Update Clear

Key:325 Created:19/6/2018 17:16:07 Last update:20/6/2018 10:51:05

Start_api_request_n_code	3814
Batch_status	(2300) Wait for human visual check in import database Import_database
Batch_status	<div><div>Remove from import database</div><div>Ready with human visual check in import database</div></div>
Batch_test	1
Metadata_path_name	e:\internet\seadatanet\CDI_V5_bestanden_aangeleverd\1526\1526_325\
Metadata_zip_file_name	demo4.zip
Metadata_files_count	15
Metadata_files_ok	15
Metadata_files_error	0
Metadata_duplicates_identifiers	0
Metadata_duplicates_batch	0
Metadata_duplicates_production	0
Metadata_duplicates_batch_log	
Metadata_duplicates_production_log	

IM: Confirmed and programmes take over again

Import Batch | Seadatanet V5 | Replication Manager | +

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110% | Search










Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Key -	Input_date -	Last_update -	Batch_test -	Batch_status -	Metadata_files_count -	Metadata_files_ok -	N_code -	Edit
325	19/6/2018 17:16 PM	20/6/2018 11:05 AM	True	(2320) Ready with human visual check in import database	15	15	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

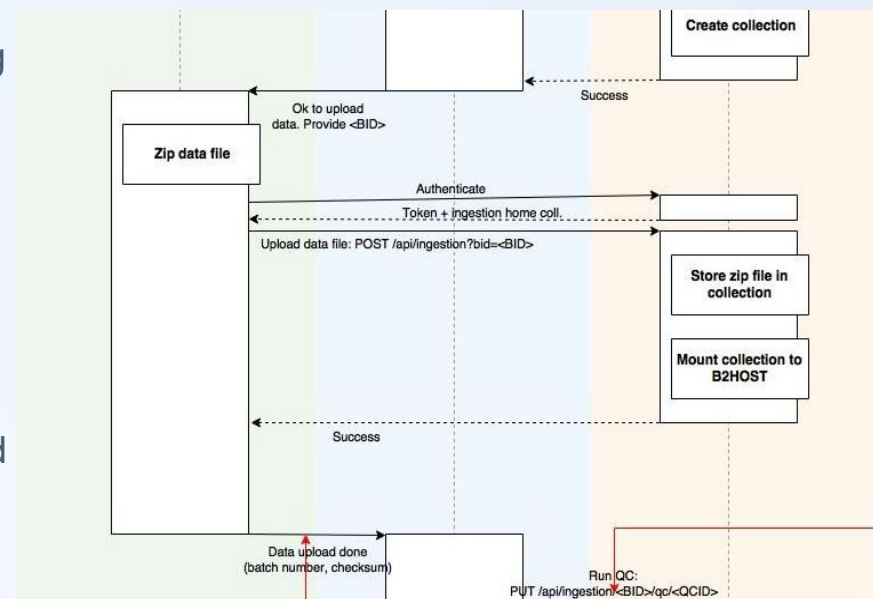
Step 4: Uploading a databatch

The problem

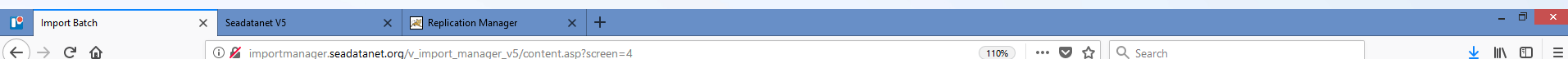
IM requests to RM to upload specified datafiles (matching CDI's with datafiles). RM has one zip file with data (!) to upload

The solution




1. RM holds files in a folder
2. RM receives from IM a Batch ID (BID), CDI_identifiers, and formats to use for upload
3. IM creates BID and creates the virtual folder on cloud
4. RM uploads the zip file to the folder in cloud
5. HTTP API confirms upload completed to RM
6. RM notifies Import Manager that upload is ready.



IM: Import manager triggers RM to upload data












Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch   

IMPORT BATCH

Free search

SEARCH CLEAR

Key -	Input_date -	Last_update -	Batch_test -	Batch_status -	Metadata_files_count -	Metadata_files_ok -	N_code -	Edit
325	19/6/2018 17:16 PM	20/6/2018 11:06 AM	True	(3000) Wait for uploading datafiles from partner	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

RM: Datafiles submitted to cloud

Import Batch

Seadatanet V5

Replication Manager

83.163.127.252:8080/rm/BatchesHistory

Search

<input type="checkbox"/>	2018-06-14 16:59:48	demo12.zip
<input type="checkbox"/>	2018-06-14 16:59:52	demo13.zip
<input type="checkbox"/>	2018-06-14 16:59:55	demo14.zip
<input type="checkbox"/>	2018-06-14 17:00:00	demo15.zip
<input type="checkbox"/>	2018-06-14 16:59:09	demo6.zip
<input type="checkbox"/>	2018-06-14 16:59:24	demo7.zip
<input type="checkbox"/>	2018-06-14 16:59:28	demo8.zip
<input type="checkbox"/>	2018-06-14 16:59:33	demo9.zip

Current batch

name	batch_number	request_id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
demo4	325	939cb7b3-5da5-46c6-95a5-dbed02ef610d	2018-06-20 10:35:35	[INGESTION_PENDING] Ingestion started - in process	[ARCHIVED] CDIs batch harvested and archived	[WAITING_PID] waiting for PIDs	demo4 (15 files)

Batches in queue

name	batch_number	request_id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
demo5			2018-06-20 10:35:35	[IN_QUEUE_READY] In queue, CDIs and data ready in queue	[CHECK_READY] Check is done	[IN_QUEUE_ZONE] Moved to Data queue directory	demo5 (15 files)

Batches cancelled

name	batch_number	request_id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
------	--------------	------------	-----------	---------------------	-------------------	-------------------	------------

IM: Datafiles on cloud, now ready to unzip and moved – Docker application on cloud

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 19 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
325	19/6/2018 17:16 PM	20/6/2018 11:08 AM	True	(3100) Unpack datafiles	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	
310	6/6/2018 12:34 PM	7/6/2018 10:20 AM	True	(2180) Removed from import database	14	14	-	

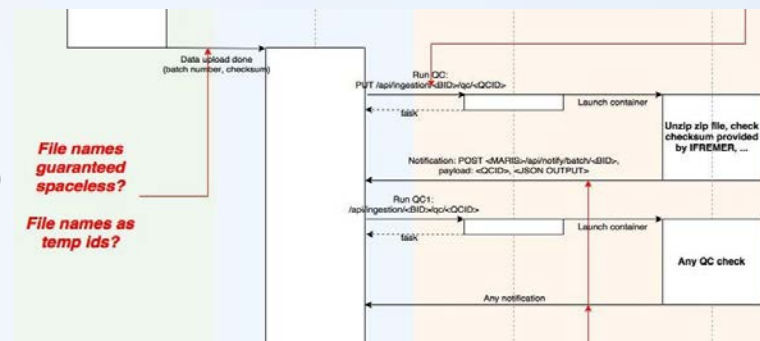
Step 5: Unzip and run datachecks

The problem

QC developer wants to run an action (unzip/quality check) on every uploaded batch

The solution

1. QC Developer writes a program to test each file of a specific folder
 1. QC developers pushes the container for this binary to EUDAT HUB
 2. QC identifier available in the end (QCID)
2. Import Manager calls PUT on /api/ingestion/BID/qc/QCID
3. HTTP API runs an asynchronous task T for QCID mounting BID folder
4. T ends and calls POST on MARIS API with:
 1. The JSON output from the task/JAVA code incl QCID
5. Next action



Technical QC implementation status

- Checks implemented:
 - Compare checksum of file
 - Compare length of file
 - Unzip file
 - Compare nr of files
 - Check length of files > 0
 - Check for each CDI_identifier if requested files and formats exist
- All errors reported back to IM, logs are user and shown in CMS.
- Erroneous records removed from batch, but batch moves forward as long as ok records exist.
- No more records, or batch cancelled by data manager, start again.

IM: Unpack datafiles done, plus technical checks (number of files, file names, etc)

importmanager/v_import_manager_v5/content.asp?screen=4

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

IMPORT BATCH

Free search:

SEARCH CLEAR

Found 20 Show (1-10) < >

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
326	20/6/2018 12:28 PM	20/6/2018 13:26 PM	True	(3120) Unpack datafiles ready	15	14	-	
325	19/6/2018 17:16 PM	20/6/2018 12:28 PM	True	(2180) Removed from import database	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

Now a sequence of additional QC checks will run

- ODV quality checks
- Conversion to other formats
- Validation
- Etc.
- Facilitated by Octopus, plus additional programmes

=> Ready to integrate, in progress

IM: Unzip and checks are done, left-over files ready to be moved. RM manager should check logs and confirm (Live!)

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 20 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
326	20/6/2018 12:28 PM	20/6/2018 13:26 PM	True	(7000) Wait for human approval go to production	15	14	-	
325	19/6/2018 17:16 PM	20/6/2018 12:28 PM	True	(2180) Removed from import database	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:22 AM	11/6/2018 10:10 AM	True	(8920) Batch in production; Ready for a new batch	2	2	-	

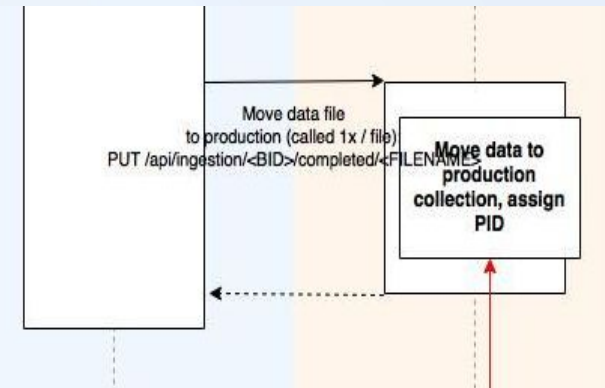
Step 6: Move validated data in production

The problem

- Import manager has a matrix of Quality checks for the batch
- All batch records that passed all quality checks is to move in production

The solution

1. IM calls HTTP API with the filename for each approved file
2. HTTP API asynchronously triggers the ingestion process of IRODS
3. iRODS moves the collection to Production
4. Automatic irules are triggered for each file to generate the PID
5. HTTP API supplies IM a JSON with filenames and their PIDs (per format)



IM: Human approval to move to production

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Import Batch

Home

Help

Refresh

IMPORT BATCH

Free search

SEARCH

CLEAR

CURRENT ACTION

STATUS OVERVIEW

METADATA LOGFILES

Results

Found 20

Show 1

Update

Clear

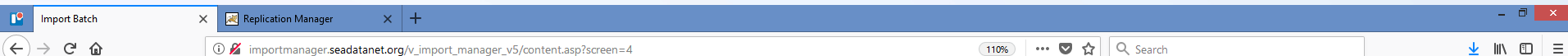
Key:326

Created:20/6/2018 12:28:53

Last update:20/6/2018 13:26:57

Start_api_request_n_code	4849
Batch_status	(7000) Wait for human approval go to production
Batch_status	<div><div>Remove from import database</div><div>Ready human approval go to production</div></div>
Batch_test	1
Metadata_path_name	e:\internet\seadatanet\CDI_V5_bestanden_aangeleverd\1526\1526_326\
Metadata_zip_file_name	demo5.zip
Metadata_files_count	15
Metadata_files_ok	14
Metadata_files_error	1
Metadata_duplicates_identifiers	0
Metadata_duplicates_batch	0
Metadata_duplicates_production	0
Metadata_duplicates_batch_log	
Metadata_duplicates_production_log	

IM: Ready to go for final steps













Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
326	20/6/2018 12:28 PM	20/6/2018 13:46 PM	True	(7020) Ready human approval go to production	15	14	-	
325	19/6/2018 17:16 PM	20/6/2018 12:28 PM	True	(2180) Removed from import database	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

IM: Extract metadata from ODV for CDI search

Import Batch Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 20 Show (1-10)

Key -	Input_date -	Last_update -	Batch_test -	Batch_status -	Metadata_files_count -	Metadata_files_ok -	N_code -	Edit
326	20/6/2018 12:28 PM	20/6/2018 13:46 PM	True	(7200) Extract metadata from ODV	15	14	-	
325	19/6/2018 17:16 PM	20/6/2018 12:28 PM	True	(2180) Removed from import database	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

IM: View of all steps sofar – always available (1 error)

SDC Import/RSM manager Import Batch Seadatanet V5 Replication Manager +

importmanager/v_import_manager_v5/content.asp?screen=4

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

Free search

SEARCH CLEAR

CURRENT ACTION	STATUS OVERVIEW	METADATA LOGFILES
2018-06-15T10:40:31	unpack_datafiles_ready	45 Data_file_count does not match unzipfiles given: 15 real:17
20180615T10:06:27	unpack_datafiles	no errors
20180615T12:38:43	upload_datafiles_ready	no errors
20180615T10:06:41	upload_datafiles	no errors
20180615T10:06:40	human_activity_metadata_ready	no errors
20180615T10:06:34	human_activity_metadata	no errors
20180615T10:06:42	front_import_database	no errors
20180615T10:06:42	front_import_database	no errors
20180615T10:06:58	Addition to import database completed	no errors
20180615T10:06:58	Add to import database	no errors
20180615T10:06:55	metadata_to_load_database_ready	no errors
20180615T10:06:55	metadata_to_load_database	no errors
20180615T10:06:41	metadata_download_ready	no errors
20180615T12:21:50	metadata_download	no errors

Results Found 16 Show 1 Update Clear

IM: Extraction ready (later add consistency between ODV and CDI checks)

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 20 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
326	20/6/2018 12:28 PM	20/6/2018 13:47 PM	True	(7220) Extract metadata from ODV ready	15	14	-	
325	19/6/2018 17:16 PM	20/6/2018 12:28 PM	True	(2180) Removed from import database	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

IM: Ready to add to cloud production - retrieve PID's!

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 20 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
326	20/6/2018 12:28 PM	20/6/2018 13:47 PM	True	(8200) Add to production database	15	14	-	
325	19/6/2018 17:16 PM	20/6/2018 12:28 PM	True	(2180) Removed from import database	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

IM: PID's received from cloud - and move CDI's to production

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 20 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
326	20/6/2018 12:28 PM	20/6/2018 13:47 PM	True	(8220) waiting for addition to production database completed	15	14	-	
325	19/6/2018 17:16 PM	20/6/2018 12:28 PM	True	(2180) Removed from import database	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

IM: PID's sent to RM, version assigned, archived

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 20 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
326	20/6/2018 12:28 PM	20/6/2018 13:48 PM	True	(8300) PID's send to Replication Manager	15	14	-	
325	19/6/2018 17:16 PM	20/6/2018 12:28 PM	True	(2180) Removed from import database	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

RM: PID's sent to RM, version assigned, archived

Import Batch

Replication Manager

83.163.127.252:8080/rm/BatchesHistory

...

Search

2018-06-14 16:59:48

demo12.zip

2018-06-14 16:59:52

demo13.zip

2018-06-14 16:59:55

demo14.zip

2018-06-14 17:00:00

demo15.zip

2018-06-14 16:59:09

demo6.zip

2018-06-14 16:59:24

demo7.zip

2018-06-14 16:59:28

demo8.zip

2018-06-14 16:59:33

demo9.zip

Current batch

name	batch_number	request_id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
demo5	326	51e64333-ada2-49af-990d-8228318be323	2018-06-20 13:45:23	[PRODUCTION] In production	[ARCHIVED] CDIs batch harvested and archived	[VERSIONED] archived with version number	demo5 (15 files)

Batches in queue

name	batch_number	request_id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
------	--------------	------------	-----------	---------------------	-------------------	-------------------	------------

Batches cancelled

name	batch_number	request_id	sent date	Batch global status	Batch CDIs status	Batch Data status	CDIs files
demo4	325	939cb7b3-5da5-46c6-95a5-dbed02ef610d	2018-06-20 11:05:39	[CANCELLED_BY_IM] Cancelled by the IM	[ARCHIVED] CDIs batch harvested and	[WAITING_PID] waiting for PIDs	demo4 (15

IM: CDI's loaded in production database

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 20 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
326	20/6/2018 12:28 PM	20/6/2018 13:50 PM	True	(8500) Add to production database	15	14	-	
325	19/6/2018 17:16 PM	20/6/2018 12:28 PM	True	(2180) Removed from import database	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

IM: CDI's loading to production database complete

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 20 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
326	20/6/2018 12:28 PM	20/6/2018 13:50 PM	True	(8520) Addition to production database complete	15	14	-	
325	19/6/2018 17:16 PM	20/6/2018 12:28 PM	True	(2180) Removed from import database	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

IM: Database loaded, now going to create search tables, elastic indexes, layers, etc – once per day at night – this takes some hours

Import Batch

Replication Manager

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 20 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
326	20/6/2018 12:28 PM	20/6/2018 13:52 PM	True	(8600) Add to production web interface	15	14	-	
325	19/6/2018 17:16 PM	20/6/2018 12:28 PM	True	(2180) Removed from import database	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

IM: Done, ready for a last message to RM!

Import Batch

Replication Manager

Seadatanet V5

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

IMPORT BATCH

Free search

SEARCH CLEAR

Found 20 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
326	20/6/2018 12:28 PM	20/6/2018 13:57 PM	True	(8620) Addition to production web interface complete	15	14	-	
325	19/6/2018 17:16 PM	20/6/2018 12:28 PM	True	(2180) Removed from import database	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	
311	7/6/2018 10:32 AM	11/6/2018 10:19 AM	True	(8920) Batch in production; Ready for a new batch	3	3	-	

IM: Success! Ready for a new batch! (if next one in queue it automatically continues)

Import Batch

Replication Manager

Seadatanet V5

importmanager.seadatanet.org/v_import_manager_v5/content.asp?screen=4

110%

Search

Login code (1526) Rijkswaterstaat Water, Traffic and Environment

Import Batch

Free search

SEARCH CLEAR

Found 20 Show (1-10)

Key	Input_date	Last_update	Batch_test	Batch_status	Metadata_files_count	Metadata_files_ok	N_code	Edit
326	20/6/2018 12:28 PM	20/6/2018 14:00 PM	True	(8920) Batch in production; Ready for a new batch	15	14	-	
325	19/6/2018 17:16 PM	20/6/2018 12:28 PM	True	(2180) Removed from import database	15	14	-	
324	19/6/2018 15:18 PM	19/6/2018 15:37 PM	True	(8920) Batch in production; Ready for a new batch	15	13	-	
323	15/6/2018 13:33 PM	19/6/2018 15:12 PM	True	(2180) Removed from import database	15	15	-	
322	15/6/2018 12:22 PM	15/6/2018 13:06 PM	True	(8920) Batch in production; Ready for a new batch	15	15	-	
317	11/6/2018 13:21 PM	11/6/2018 13:57 PM	True	(8920) Batch in production; Ready for a new batch	100	80	-	
316	11/6/2018 11:55 AM	11/6/2018 12:38 PM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
314	11/6/2018 11:37 AM	11/6/2018 11:53 AM	True	(8920) Batch in production; Ready for a new batch	2	1	-	
313	11/6/2018 10:20 AM	11/6/2018 11:36 AM	True	(8920) Batch in production; Ready for a new batch	26	23	-	

Batch active in interface (example record – in test)


Import Batch

Replication Manager

Seadatanet V5

cdi_front_test/v_cdi_v5/welcome.php

Search


PAN-EUROPEAN INFRASTRUCTURE FOR
OCEAN & MARINE DATA MANAGEMENT

[LOGIN](#)
[DATASET BASKET 0](#)

[NEW SEARCH](#)
[SEARCH RESULTS](#)

[SUMMARY](#)

DETAILS

Results | show (1 - 1) | Prev | Next

WHAT?

Data set name

Discipline

Parameter groups

Discovery parameter

P01

GEMET-INSPIRE themes

Abstract

Data format

Data set creation date

WHERE?

Map

Measurements at Bath in 1955 by RIKZMON_TEMP using THERMMTR

Physical oceanography

Water column temperature and salinity

Temperature of the water column

Julian Date (chronological)


Temperature of the water body by in-situ thermometer

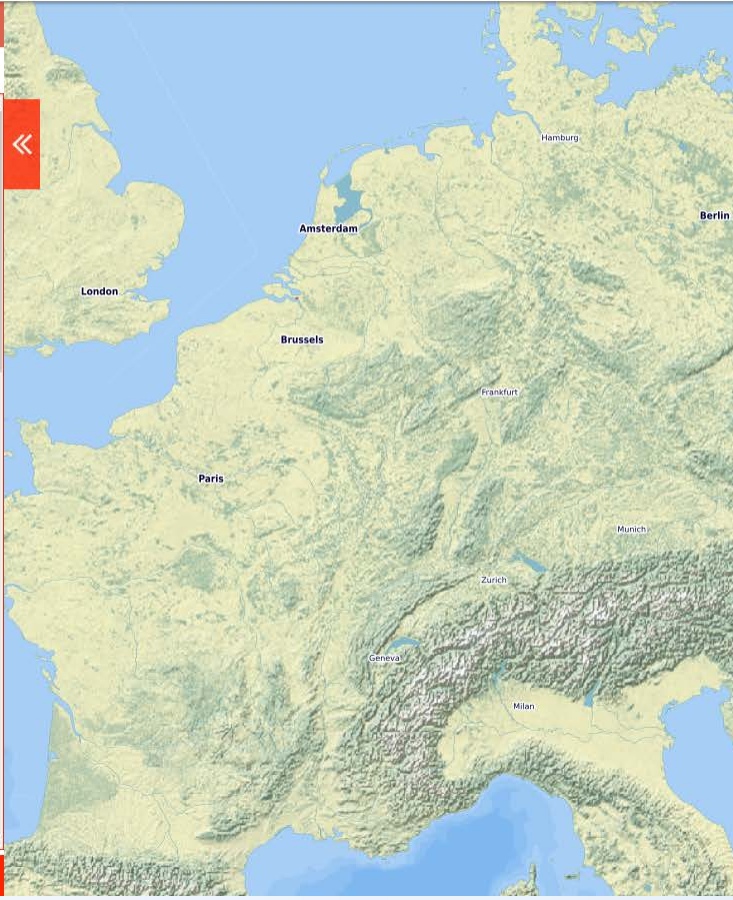
Oceanographic geographical features

BATH-1955

Ocean Data View ASCII input **Version 0.3**

20140306





ADD TO BASKET

XML

PRINT

Restricted data (in development)

- CDI metadata same flow
- Data files extracted at same time as metadata. Copy stored (for each version)
- Files will be temporarily released to cloud directory of user upon request RSM

Next steps

- Finalise integration tests and restricted data flow
- Focus on errors – what if a system is down, a service does not respond, crashes etc
- Fool (or full) proof version
- Create documentation
- Planning:
 - See implementation plan

- END -