



SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

Interoperability with other communities

Driven by a.o. the project Genesi-DEC: Establishing interoperability with the SDN Infrastructure

Peter Thijsse (MARIS) – SDN plenary meetin Rhodes September 2012

Content

- Why working on interoperability of CDI?
- Defined interoperability alternatives/use cases
- Solutions & developments towards interoperability using:
 - OpenSearch - status
 - OpenID extensions - status
 - Machine2machine interface for RSM
 - Webservices developed
 - Use for external portals
 - Use for internal data buffering for special users
- Planning

1. Why working on interoperability of CDI?

Why interoperability

- CDI system has developed into a system containing a large volume of data
- System must be prepared to serve data to external communities (Wise Marine, EEA, MyOcean, EuroGeoss, Genesi-DEC, etc)
- Current human users also request automatic application access
- Benefit of central CDI solution: strong foundation of connected data centers => Now extra services developed on top of this foundation. Underlying data centres do not have to develop similar additional services.

Interoperability

In order to make CDI (meta)data available to overarching portal systems or special users and applications, interoperability solutions are needed for:

- Discovery
- Security
- Ordering/downloading data

Several solutions have been defined with increased level of complexity and for different users.

2. Defined interoperability phases and solutions

Alternative 1 – Discovery via metadata

- CDI metadata made accessible for external portals, possibly in an aggregated way
 - Enabling users in external portals to oversee the combined offer.
 - However, for performing the actual data access request and downloading users are routed to the respective portals and their services.
- => Solution 1: Achieved **via the GI-CAT model** supplying aggregated metadata records that has been set-up by SDN to interoperate by CS-W with INSPIRE/GEOSS
- => Solution 2: **Via OpenSearch** protocol searching on registered “virtual aggregated series” (SDN data e.g. divided per discipline per partner)

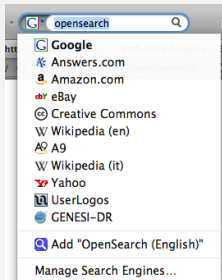
Discovery via OpenSearch

- OpenSearch communication layer installed on top of CDI database
- Access point per aggregation (RDF):
 - Per discipline
 - Per parameter/instrument/partner/..
- Prototype running for GeoSeas collection

Quick recap of OpenSearch

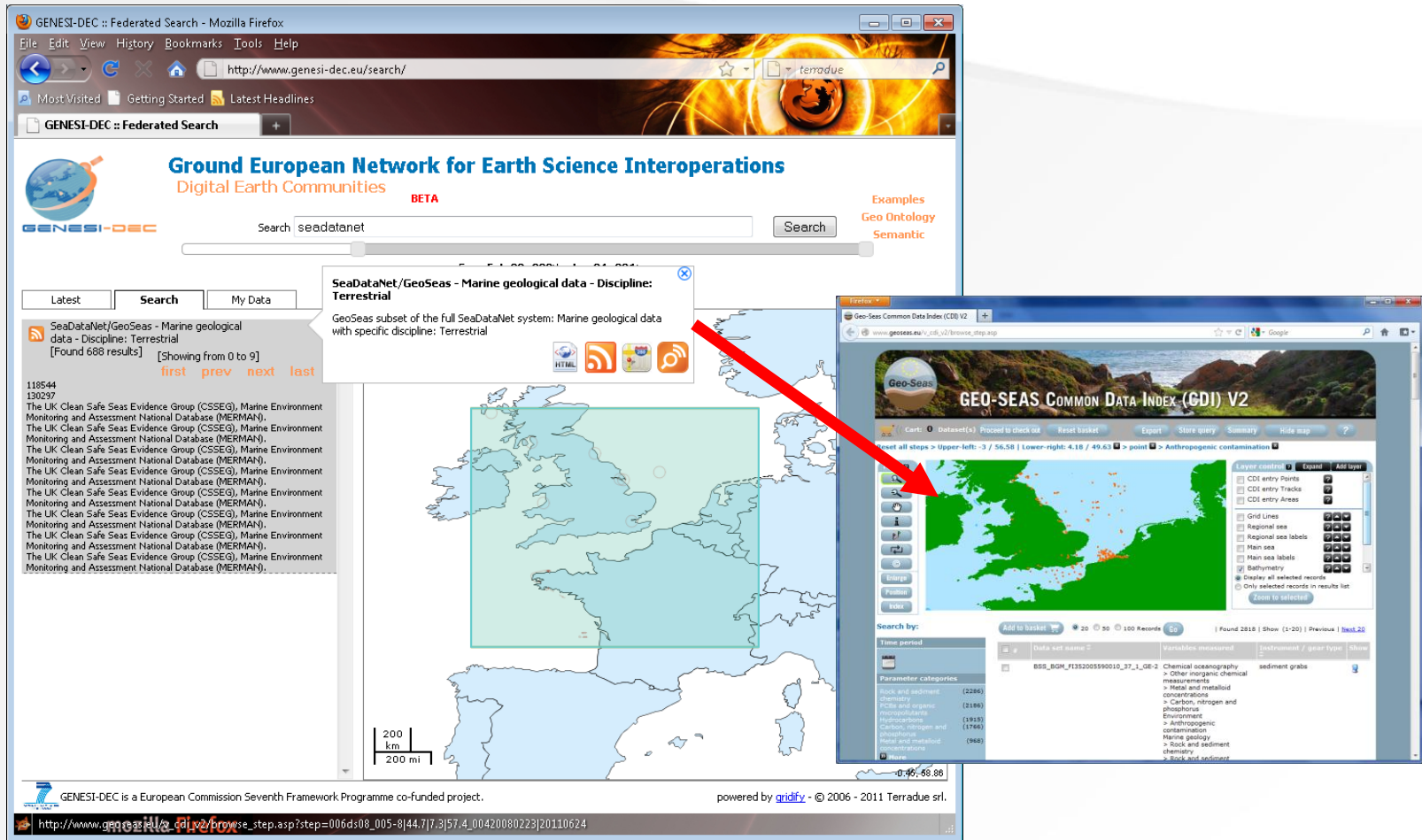
- Provides metadata about the contents along with a set of URL Templates which illustrate the parameters accepted by the service and the variety of output formats in which results can be obtained.
- The OpenSearch request interface is **simple**, consisting of a description of a HTTP GET request with a series of optional key-value parameters that can be used to constrain the search:
 - Free search
 - Geospatial (area or point+radius)
 - Temporal (from to)
 - So.. No vocabs, no special search fields..

```
<Url type="text/html"  
template="http://example.com/?q={searchTerms?}&pw={startPage?}"/>
```



```
http://www.google.com/?q=question  
http://www.google.com/?q={searchTerms}
```

Implementation: Discovery from Genesi-DEC to SDN



GENESI-DEC :: Federated Search - Mozilla Firefox
 http://www.genesi-dec.eu/search/
 Search: seadatanet
 Results: SeaDataNet/GeoSeas - Marine geological data - Discipline: Terrestrial
 [Found 688 results] [Showing from 0 to 9]

Geo-Seas Common Data Index (CDI) V2
 Search by: [Time period] [Parameter categories] [Data set name] [Variables measured] [Instrument / gear type] [Show]
 Found 2818 | Show (1-20) | Previous | Next

Alternative 2 - Adding security

Same as alternative 1, but adding interoperability between the authentication / authorization services :

Support SSO that users have to register only once (in investigated case either SDN or GENESI-DEC)

- Users can use their received personal login details (and associated user roles) in both infrastructures
- Remaining logged in when moving from portal to repository (ideal but complex in practice!)

⇒ Solved by adding OpenID extension to SDN CAS service.

Security interoperability using OpenID protocol (Alt 2)

- OpenID: Is an identity provider, that can be requested in a standard way to verify your account and confirm this.
- Genesi-DEC (or other OpenID using community) users will be able to use their account in SeaDataNet.
- Extend the SDN CAS and registration pages:
 - Allow registration by Genesi-DEC OpenID users
 - Pre-fill SDN user registration page using OpenID information
 - When accepted (normal procedure) the OpenID of the user is its SDN account
 - Next login the SDN AAA service requests login at OpenID provider for this user
- At the same time Genesi-DEC will allow SDN accounts by making connection to the CAS service to check credentials.

Alternative 3 – “Full” interoperability to SDN CDI

- A machine/application handles for the user the full process from query, to order, to request access, to trace progress and if ok, to download data sets
- Requires interoperability solutions at all levels:
 - metadata access at original granularity
 - Solution for authentication / authorization systems
 - and **machine interaction to SDN RSM.**

3a. Central component: Machine-to-machine (M2M) services for the CDI RSM

In short:

- All steps in the CDI RSM should be able to be triggered and received by a machine application (for a known SDN user):
 - Posting orders (as a query) to SeaDataNet CDI via the CDI RSM
 - Allow status request (max per hour) by a known application
 - Supply link to download page at data center when data order is ready
 - DM to allow download by the application directly (without another login)
- Note: RSM already has web services to communicate with DM. As part of this solution a second RSM web services has been built to communicate with an external client.
- Used as basis:
 - M2M supports external (modeling) applications that need to harvest data sources regularly
 - M2M services can also be used internally for buffering data for special users
 - External portal can use the M2M services after discovery for ordering and downloading data by their users. E.g. Genesi-DEC as example of a portal for discovery where data can be ordered and downloaded from various repositories.

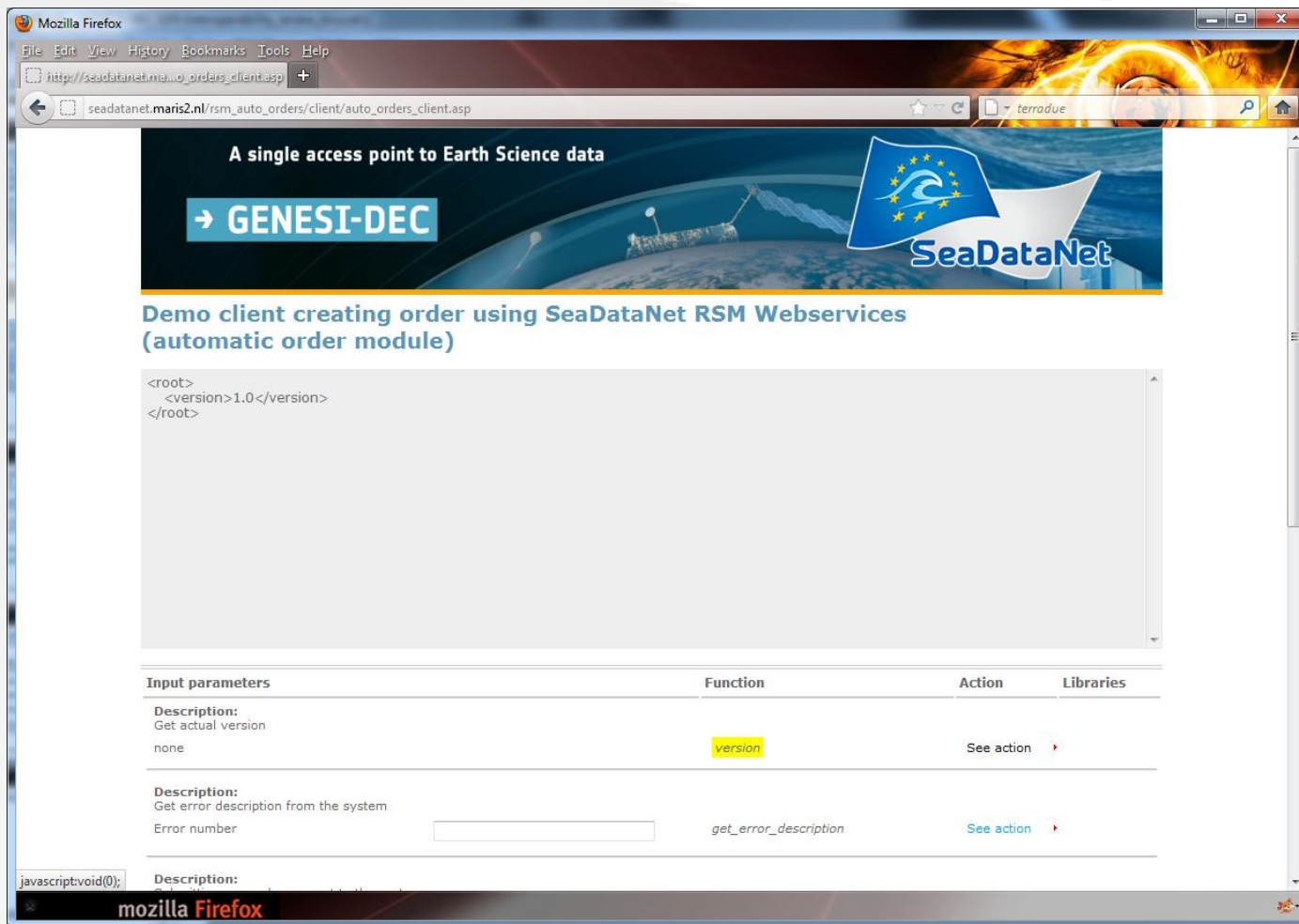
Machine2machine services – Technical specs

- SOAP web services
- Based on functions of the SeaDataNet RSM (Request Status Manager)
- Security IP recognition from trusted portal, plus user recognition via SDN AAA services (OpenID support needed to allow Genesi-DEC users!)
- Client interface available for demonstration
- Ready to be integrated
- Available via:
 - WSDL:
http://seadatanet.maris2.nl/rsm_auto_orders/auto_orders.asmx?WSDL
 - Or demo client:
http://seadatanet.maris2.nl/rsm_auto_orders/client/auto_orders_client.asp

Machine2machine services – Functions

- version: Get actual version. (can also be used for test if service is active)
- get_error_description: Get error description from the system, translating error codes.
- submit_query_order: Submitting any order request to the system.
- get_order_info: Get information of all not completed orders.
- get_order_status: Get status information from one order.
- get_order_downloads: Get download information from one order including download paths to zipfile at datacenter.

Machine2machine services – Client preview



A single access point to Earth Science data

→ GENESI-DEC

SeaDataNet

Demo client creating order using SeaDataNet RSM Webservices (automatic order module)

```
<root>
<version>1.0</version>
</root>
```

Input parameters	Function	Action	Libraries
<p>Description: Get actual version</p> <p>none</p>	version	See action	
<p>Description: Get error description from the system</p> <p>Error number <input type="text"/></p>	get_error_description	See action	

javascript:void(0);

mozilla Firefox

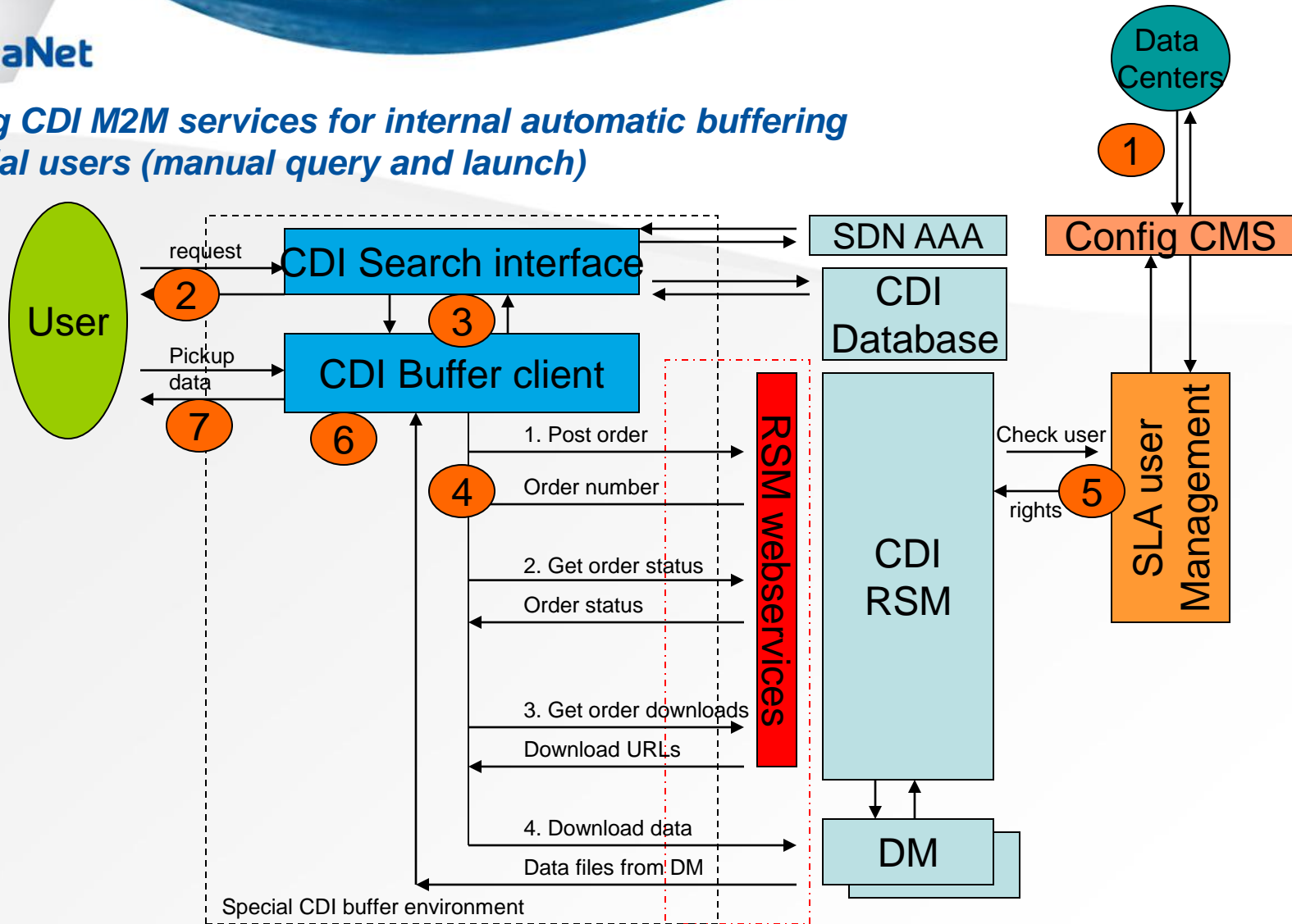
3b. Using CDI M2M services for internal automatic buffering for special users (Manually triggered order!)

1. **Configure:** A special user (e.g. MyOcean) makes an agreement with (some of) the datacenters of SeaDataNet. Via the config CMS of the user management the data centers set the special rights of this user for their data.
2. **Login:** User logs in to special CDI interface
3. **Create order:** Via CDI-Buffer interface user posts order as a query
4. **Collect:** From that moment a SDN buffer client to CDI M2M services handles the complete order automatically until download of the released data from the DMs
5. **Check rights:** During process the RSM checks special rights of user in SLA user management, and if necessary the SLA can “override” normal access restrictions
6. **Store:** Datasets are placed on the central CDI server on a special user page (AAA login needed). User page shows:
 - Order number and date
 - Query details on request
 - Order status (what percentage ready for download/denied/pending)
 - Files immediately ready for download collected in one zipfile
7. **Download:** User downloads 1 file per order



SeaDataNet

3b. Using CDI M2M services for internal automatic buffering for special users (manual query and launch)





SeaDataNet

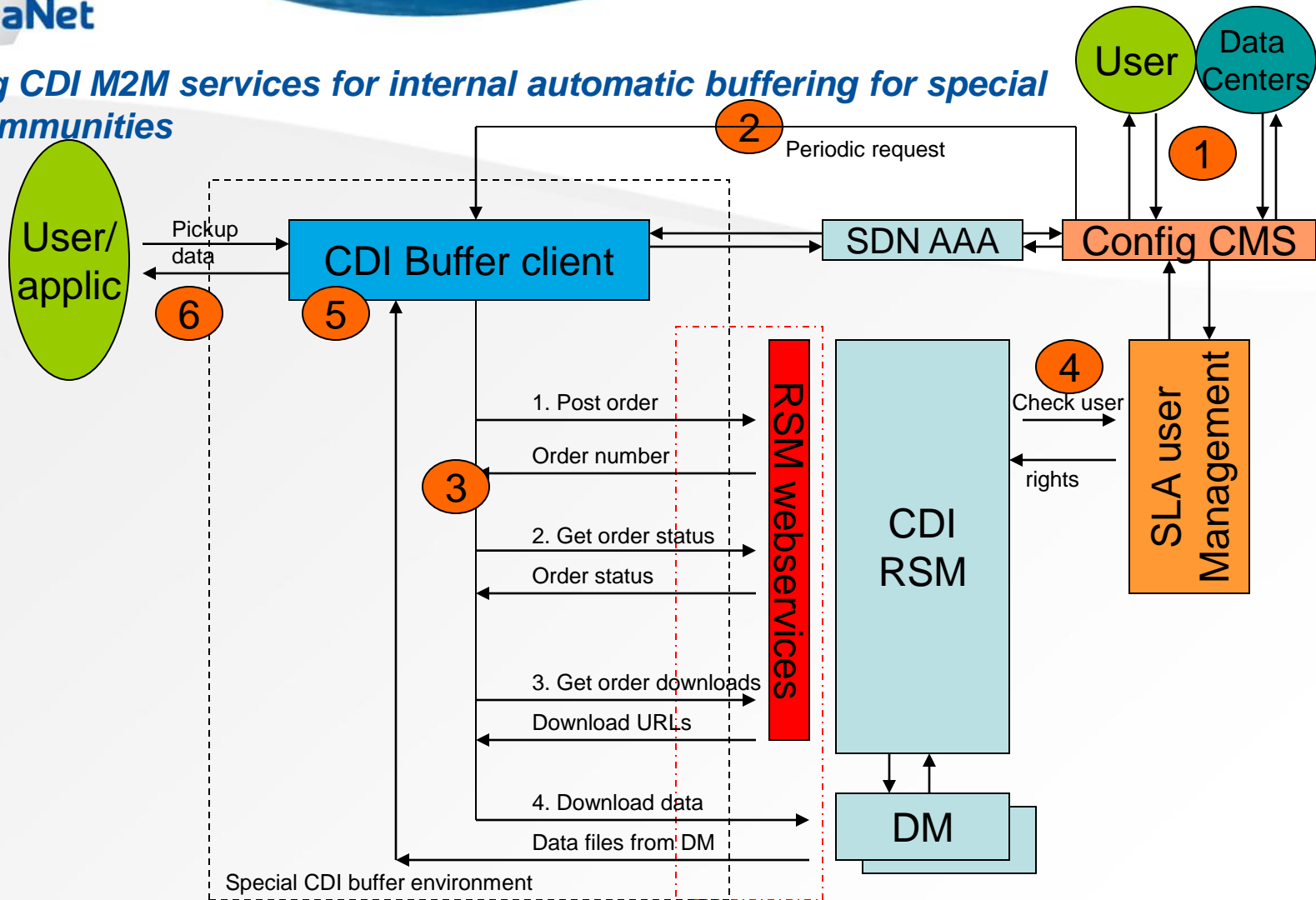
3c. Using CDI M2M services for internal automatic buffering for special communities (periodic automatic action!)

1. **Configure:** A special user (e.g. EEA) makes an agreement with (some of) the datacenters of SeaDataNet. Via the config CMS of the user management the following settings are made by data center and user:
 - Datacenter sets the special rights of this user for their data
 - User (or SDN user desk) specifies the data order that this user needs and the updating frequency of the buffer
2. **Start order:** At the required time and interval the config service posts a query order to the CDI-Buffer service
3. **Collect:** From that moment a SDN client to CDI M2M services handles the complete order automatically until download of the released data from the DMs
4. **Check rights:** During process the RSM checks special rights of user in SLA user management, and if necessary can “override” normal access restrictions
5. **Store:** Datasets are placed on the central CDI server on a special user page (AAA login needed). User page shows:
 - Order number and date
 - Query details on request
 - Order status (what percentage ready for download/denied/pending)
 - Files immediately ready for download collected in one zipfile
6. **Pick up data:** User or user application downloads latest version of data collection



SeaDataNet

3c. Using CDI M2M services for internal automatic buffering for special users/communities



3c. Full interoperability from external portal using Machine-to-machine interfaces

(Currently developed as a prototype only!)

- Use of M2M services by a portal to present order and download services for their users. E.g.: Genesi-DEC as example of a portal for discovery where data can be ordered and downloaded from various repositories

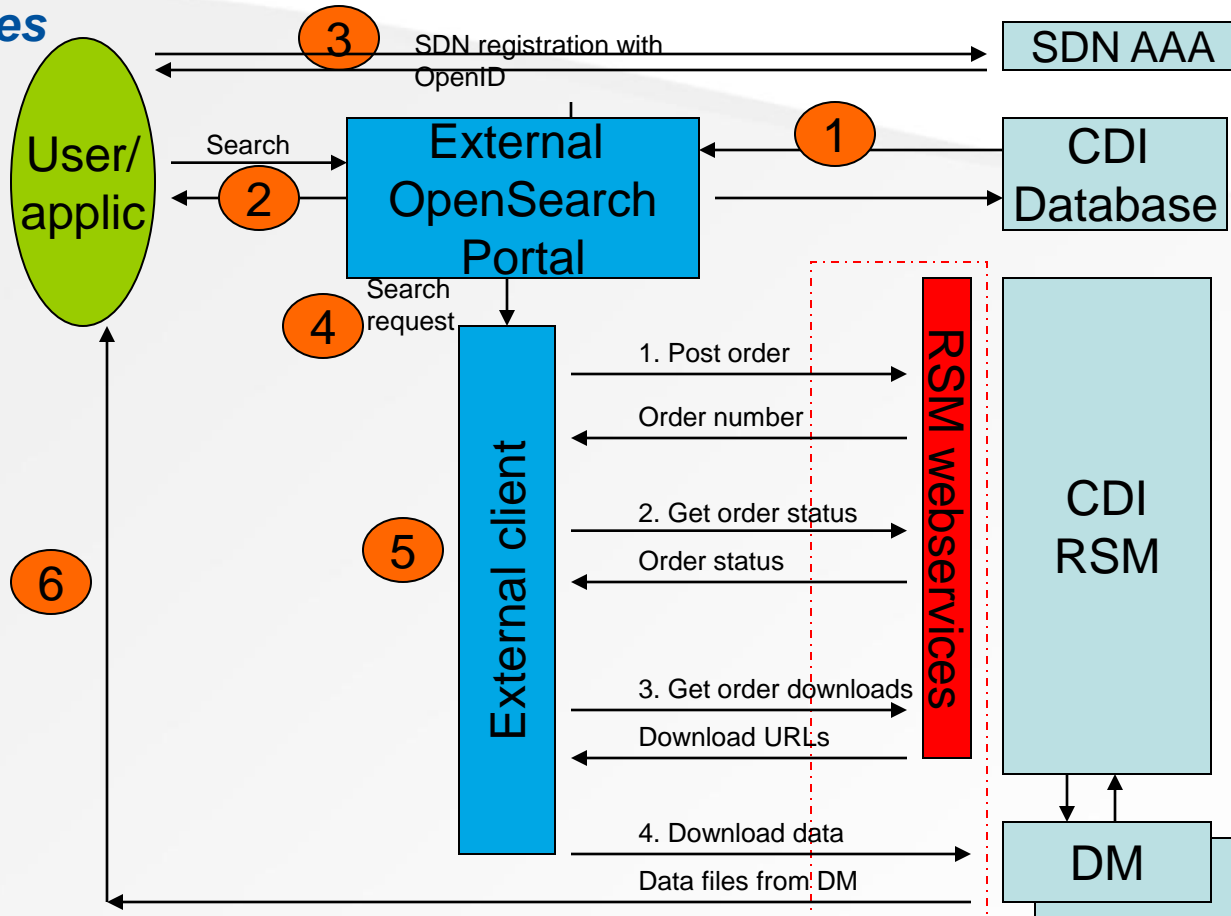
In steps:

1. Registration of OpenSearch endpoints on external portal (alternative could be GICAT)
2. Basic discovery via OpenSearch
3. External user registers as SDN user at SDN portal
4. Search result transferred to machine client at external portal
5. Machine client transforms request to order for M2M RSM service and handles all steps in process, presenting progress and eventually download links
Note: Request only allowed for trusted servers (IP recognition)
6. User pickup data directly from SDN download managers



SeaDataNet

3c. Full interoperability from external portal using Machine-to-machine interfaces



IP protected
communication

Planning

- Oct 2012: Test implementation of external client by Genesi-DEC to use (and test) CDI RSM web services
- Oct-Dec 2012: RSM web services fully tested and expanded
- Oct 2012 -Sept 2013: First full specification, then development of CDI buffer services for special users (first version Dec 2012 for MyOcean2)



PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

Thank you!