DELIVERABLE D4.3

COMSODE publication platform - Open Data Node - final

---

Project: Components Supporting the Open Data Exploitation
Acronym: COMSODE
Contract Number: FP7-ICT-611358
Start date of the project: 1st October 2013
Duration: 24 months, until 31st September 2015

Date of preparation: 1.7.2015
Author(s): Peter Hanečák, Svetozár Krchnavý, Ivan Hanzlík
Responsible of the deliverable: Peter Hanečák
Email: peter.hanecak@eea.sk
Reviewed by: Oskár Štoffán, Jan Kučera, Jakub Klimek, Tomáš Knap
Status of the Document: final version
Version: 1.0
Dissemination level: PU (Public)
Table of Contents

1. Executive summary ........................................................................................................... 3
2. Deliverable context ........................................................................................................... 4
   Purpose of the deliverable ................................................................................................. 4
   Related documents ........................................................................................................... 4
3. Methodology used ............................................................................................................ 5
   Testing .............................................................................................................................. 5
   Step 1 - development plan .............................................................................................. 6
   Step 2 - implementation of ETL & enrichment functionality ........................................... 9
   Step 3 - implementation of Publication functionality ..................................................... 9
   Step 4 - implementation of Cataloguing functionality .................................................... 10
   Step 5 - implementation of Management functionality ................................................ 10
4. Final release of Open Data Node ...................................................................................... 11
   UnifiedViews .................................................................................................................. 12
   CKAN ............................................................................................................................... 14
   Single Sign-On ............................................................................................................... 16
   Packaging ...................................................................................................................... 18
   Virtuoso (Open Source version) ..................................................................................... 19
5. Support infrastructure created for Open Data Node components ............................... 20
   Support infrastructure for Open Data Node itself ......................................................... 20
6. Services of Pilot publishing project ............................................................................... 21
7. List of ODN releases ....................................................................................................... 22
8. Future roadmap ............................................................................................................... 23
1. Executive summary

Deliverable 4.3 consists mainly of the latest stable release of COMSODE Open Data publication platform named Open Data Node (also referred to as ODN) version 1.0.3, along with the documentation sufficient to deploy, use and maintain the solution. The delivery is licensed under a set of Open Source licenses and all parts are available on the project website for free download.

Deliverable is a snapshot of what is publicly available as output of Task 4.1 at the end of WP4, i.e. M21 (June 2014, major project milestone with several deliverables due). Task started in M8 and continued through M12 (when deliverable D4.1 was submitted, documenting status of ODN at that time) until M21.

This document describes the main achievements of Task 4.1 available at the end of M21:

- latest stable release of Open Data Node, version 1.0.3
- support infrastructure created for its components
- services provided by Pilot publication project

And because some further releases of ODN are already planned, deliverable will also document the roadmap for next few months.

It also describes purpose of this deliverable and methodology used to produce above mentioned milestones.
2. Deliverable context

Purpose of the deliverable
The purpose of this deliverable is to build stable and publicly available Open Source implementation of Open Data publication platform.

This stable release is then further used:
- by COMSODE consortium, pilots and other users to publish Open Data,
- to develop Deliverable 4.4 (Tailored search application on-top of data published using ODN),
- to achieve project indicators due in Year 2 and
- to support further Exploitation.

Related documents
List of related documents from project:
- DoW, final version
- D2.1 - User requirements for the publication platform from target organizations, including the map of typical environments
- D2.3 - Architecture and design documentation for COMSODE development tasks
- D5.1 - Methodology for publishing datasets as open data
- D4.1 - COMSODE publication platform - Open Data Node - for test
- D4.4 - Tailored search application on-top of data published using ODN - final
- D6.1 - Preliminary evaluation report
- D6.2 - Final evaluation report
- D7.3 - Halftime exploitation strategy
- Periodic Progress Report, due November 2014
3. Methodology used

The final goal of Task 4.1 is a delivery of Open Source publication platform for Open Data (also referred to as Open Data Node, ODN in short) which is expected to be deployed mainly within data publishing organizations and which will provide access to the published Open Data sets to general public.

Development process is part of the Task 4.1 and started in M8 (May 2014) with inputs from WP2 (Deliverable 2.3). Task ended in M21 (June 2015).

Development was divided into 5 main steps:

1. creation of development plan
2. implementation of ETL & enrichment functionality
3. implementation of Publication functionality
4. implementation of Cataloguing functionality
5. implementation of Management functionality

At the end of the Year 1 (i.e. M12, Sep 2014), we completed steps number 1 and 2. Thus Deliverable 4.1 - test version of Open Data Node - consisted of the results from those two steps and respective test version was not feature complete.

At the end of WP4 (i.e. M21, Jun 2015) steps 3 to 5 were completed and Deliverable 4.3 - final version of Open Data Node - is considered complete: it contains all basic modules planned and implements all the basic requirements stated for publication platform. But given deviations from the plan (see end of section “List of ODN releases” on page 22) further ODN releases with additional features are still being developed.

It is necessary to point out, that development process of ODN is and has to be agile. Even though we have well defined requirements, user stories, architecture and design of the platform along with plan for implementation, we have to take into account requirements and feedback from the User board, pilots and especially our currently most important implementation of the ODN as Slovak nationwide open data platform (project eDemokracia). So it is obvious that deviations from the plan and initial specifications occurred.

In the Year 2, the thorough, ongoing and well managed testing of ODN started (task of WP6). Testing provides (until M22, end of WP6, but even beyond) feedback regarding completeness and quality of implementation.

Testing

As part of WP6 (mainly tasks 6.1 and 6.2) testing of ODN started in Year 2.

Test plan for each ODN release is outlined and executed. Test results were continually reported back to WP4.
Final Evaluation Report (D6.2) also includes test result for one particular ODN release. But given that testing of one release takes approximately two weeks, there is small mismatch between D4.3 and D6.2 based on what was available at the end of M21:

- D4.3 describes ODN 1.0.3: this is the latest stable ODN release available
- D6.2 reports on ODN 1.0.2: this is the latest complete test report available

As 1.0.3 is just a bug fix release for 1.0.2 with no added features, the discrepancy between D4.3 (ODN 1.0.3) and D6.2 (ODN 1.0.2) is small and can be summed as “more pessimistic view of ODN”, given that D6.2 reports more bugs: bugs identified on release 1.0.2, many of those already fixed in release 1.0.3. This discrepancy will be addressed by Testing Report (D6.3) due next month.

**Step 1 - development plan**

Based on inputs from WP2 - for example high-level architecture of Open Data Node (Fig. 2) and effort estimates for each particular ODN module - it was decided to align the development roughly along the data flow paths: from the raw data and metadata at the beginning to the final Open Data (datasets and their metadata) at the end. Processes and steps described in Deliverable 5.1 (generic Open Data publishing methodology) follow similar pattern thus further reinforcing this decision.

So in Task 4.1, we’ve developed Open Data Node roughly from “the left” to “the right” (Fig. 1):

*Figure 1: Basic Open Data Node use-case, with flow of data illustrated*
Thus, the order of implementation was decided to be:

1. module ODN/UnifiedViews: ETL & enrichment functionality
2. module ODN/Storage: storage of results from ODN/UnifiedViews for later use in ODN/Publication
3. module ODN/Publication: functionality for publication of data dumps and APIs
4. module ODN/InternalCatalog: internal cataloguing functionality needed mainly for first phases of COMSODE Methodology (D5.1)
5. module ODN/PublicCatalog: public cataloguing functionality needed mainly for final phases of COMSODE Methodology (D5.1)
6. module ODN/Management: management and integration functionality for Open Data Node as a whole

As some modules are more closely related to each other, final plan consisted of 4 further steps described in subsequent chapters. Each step was estimated to take approx. 3 to 4 months to complete and steps were ordered one after another with small (one or two weeks) overlaps. Figure 3 illustrates the roadmap:
Figure 3: Open Data Node development road map, showing also relation to Task 4.3, Task 6.5, inputs from other Work Packages and also Year 1 milestone (Deliverable 4.1)

Later, in section “List of ODN releases” (page 22) we describe also differences between this plan and actual progress.
Step 2 - implementation of ETL & enrichment functionality

In this step, ODN development started by implementing ODN/UnifiedViews and (part of) ODN/Storage and ODN/Publication modules. Following tools were identified in WP2 for reuse:

- UnifiedViews (http://unifiedviews.eu/; formerly ODCS: https://github.com/mff-uk/ODCS) as base for ODN/UnifiedViews module (originally named as ODN/ETL)
- Apache HTTP project (http://httpd.apache.org/) as one of components for ODN/Publication module, reused as-is, without need to modify it
- PostgreSQL (http://www.postgresql.org/) as one of components for ODN/Storage, reused as-is, without need to modify it
- Virtuoso - Open Source version (http://virtuoso.openlinksw.com/) as another component for ODN/Storage module, reused as-is, without need to modify it

Majority of development effort went into redesign and stabilization of UnifiedViews, which was also - as the part of this step (and with collaboration with associated partners) - rebranded from ODCS to UnifiedViews (see the note below).

Note: Some part of development effort (partially also Exploitation) related to UnifiedViews actually started in M6 (March 2014), well ahead of planned start of Step 2 (M8). This was necessary for smooth transition from the student project (ODCS) into community backend, market-ready solution branded as UnifiedViews. Initial community consisting of Charles University in Prague (COMSODE consortium partner), EEA (COMSODE consortium partner) and Semantica.cz (Spin-off Company of some of original ODCS developers) was formed, governance structure created and licensing and development model put in place.

In line with that, also module ODN/ETL was renamed to ODN/UnifiedViews so as to signify that this module is tasked not just with ETL operations but also with data enhancement and enrichment functions, as described in DoW.

Step 3 - implementation of Publication functionality

ODN development continued in this step mainly with implementation of ODN/Publication module, with further enhancements to ODN/UnifiedViews and ODN/Storage.

WP2 identified tools for reuse, for example:

- restSQL (http://restsql.org/) as base for implementation of REST API for datasets published via Open Data Node, non-trivial development effort dedicated to enhance the tool to meet the needs of COMSODE project
- Virtuoso - Open Source version (http://virtuoso.openlinksw.com/) as provider of SPARQL endpoint functionality, to be reused as-is, without need to modify it

In reality, the actual development jumped from step 3 to step 4 after completing functionality required for publication of data dumps and SPARQL endpoint, without implementing REST API. REST API was finished later during step 4, see next section.

During Year 2, development plan was updated based on agreement with CUNI, that CUNI will contribute Payola visualisation tool (http://payola.cz/) into ODN. As of June 2015:

- Payola was rebranded to LDVMi (http://ldvm.github.io/LDVMi)
- LDVMi is being integrated into ODN (see section “Future roadmap”, page 22)
Step 4 - implementation of Cataloguing functionality

ODN development in this step consisted mainly of implementation of two modules: ODN/InternalCatalog and ODN/PublicCatalog (in D4.1 we referred to it as ODN/Catalog). Implementation was merged into one step because WP2 identified same tools to be reused for both and also because both modules are closely related through data publication use-cases.

WP2 identified following tool for reuse in both modules:
- CKAN (http://ckan.org/) as base for cataloguing functionality in Open Data Node

Non-trivial development effort was needed to enhance the tool to meet the needs of COMSODE project. For example:
- We needed to redo Debian packaging in order to optimize dependencies (SOLR and Java in particular), streamline installation and upgrades together with the rest of ODN and allow concurrent deployment of two CKAN instances on same server.
- We integrated together ODN/InternalCatalog and ODN/UnifiedViews so that data publisher does not have to jump between two tools for frequent publication activities.

In each module (or role: internal catalog, public catalog), CKAN requires different configuration and different set and amount of customization, majority of it in ODN/InternalCatalog.

While investigating what to re-use and how to implement our further customizations, we've took a closer look at data.gov.uk: This portal has several additional features implemented (compared to official CKAN), some of them of interest to COMSODE and its source code is available for re-use. But after closer investigation we found out that all such features are mixed and implemented in just one extension. So we had to reconsider our re-use plans (it was not possible for us to reuse pieces from data.gov.uk) and it also strengthened our decision regarding design and implementation of our extensions: all of our customizations are implemented as CKAN extensions, with extensions being kept separate (one customization per one extension). This eases our long maintenance burden, makes ODN more flexible and also makes it easier to re-use smaller pieces from COMSODE by others.

As mentioned in previous section, in this step we also finished implementation of REST API: given availability of DataStore functionality in CKAN (which become part of CKAN in version 2.2 and by M17 was considered sufficient for ODN) we’ve reconsidered use of restSQL and finished the implementation of REST API for Publication module re-using CKAN’s DataStore functionality.

Step 5 - implementation of Management functionality

ODN development in this step consisted mainly of implementation of module ODN/Management. WP2 has not identified any major tool for reuse in this modules as majority of the work was expected to be customization and integration between tools reused in other modules.
Later on, as the requirements and (more importantly) strengths and weaknesses of already re-used components become clearer, following tools were identified for re-use:

- midPoint (https://www.evolveum.com/midpoint/) as base for user management and Single-Sign-On (SSO)
- CAS (https://www.apereo.org/projects/cas) as the base for SSO integration for other modules (UnifiedViews, CKAN, etc.)

This gave ODN users SSO - very important feature from the user experience perspective: thanks to Single-Sign-On they have just one login and password for whole ODN (all of its modules) and they need to login or logout only once, no matter in which module.

For more details, see section “Single Sign-on” on page 22.

4. Final release of Open Data Node

Final version of Open Data Node consists of modules:

1. ODN/UnifiedViews implemented using UnifiedViews
2. ODN/InternalCatalog implemented using CKAN
3. ODN/PublicCatalog implemented (again) using CKAN and
4. ODN/Management implemented using mainly midPoint

This is supplemented with following two modules:

1. ODN/Storage implemented using general file system, PostgreSQL and Virtuoso (Open Source version)
2. ODN/Publication implementation based mainly on Apache HTTPD, which:
   a. unifies together various publicly visible components of ODN: ODN/PublicCatalog (cataloguing functionality, REST API, etc.), Virtuoso’s SPARQL endpoint and other components (like upcoming LDVMi)
   b. makes the ODN more secure by hiding internal services and exposing only public services, enforcing also encryption when authentication is needed.

<table>
<thead>
<tr>
<th>component</th>
<th>home page</th>
<th>license</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAS</td>
<td><a href="https://www.apereo.org/projects/cas">https://www.apereo.org/projects/cas</a></td>
<td>APLv2</td>
</tr>
<tr>
<td>CKAN</td>
<td><a href="http://ckan.org/">http://ckan.org/</a></td>
<td>AGPLv3</td>
</tr>
<tr>
<td>midPoint</td>
<td><a href="https://www.evolveum.com/midpoint/">https://www.evolveum.com/midpoint/</a></td>
<td>APLv2</td>
</tr>
<tr>
<td>PostgreSQL</td>
<td><a href="http://www.postgresql.org/">http://www.postgresql.org/</a></td>
<td>similar to BSD and MIT</td>
</tr>
<tr>
<td>UnifiedViews</td>
<td><a href="http://unifiedviews.eu/">http://unifiedviews.eu/</a></td>
<td>GPLv3 and LGPLv3</td>
</tr>
<tr>
<td>Virtuoso Open Source</td>
<td><a href="http://virtuoso.openlinksw.com/dataspace/doc/dav/wiki/Main/">http://virtuoso.openlinksw.com/dataspace/doc/dav/wiki/Main/</a></td>
<td>GPLv2</td>
</tr>
</tbody>
</table>
More complete and up-to-date list is maintained at

Main COMSODE development effort was dedicated to UnifiedViews and CKAN and also to
integration and packaging work required to deliver unified ODN installation. So those will be
described in more details. Additional notes are included also about Virtuoso, as only certain
versions of it are supported.

**UnifiedViews**
COMSODE project contributed and released so far 17 versions of UnifiedViews, starting with
version 0.9 (in March 2014), with version 1.3.1 used in D4.1 (September 2014) and currently
at version 2.0.4 (September 2015, part of upcoming ODN 1.1.0).

UnifiedViews is available under set of Open Source licenses and can be obtained:
- as source code from https://github.com/UnifiedViews
- as pre-build Java binaries from http://maven.eea.sk/artifactory/public/eu/unifiedviews/
- as pre-build Debian Linux packages from http://packages.comsode.eu/pool/main/u/

UnifiedViews consists of following components:
- **Core**: Main part, providing management GUI and backend running ETL and
  enrichment task (called pipelines), which are constructed from individual Data
  Processing Units (DPUs).
  - https://github.com/UnifiedViews/Core
  - license: combination of GPLv3 and LGPLv3
- **Generic Plugins**: Set of common DPUs provided for Core.
  - https://github.com/UnifiedViews/Plugins
  - license: LGPLv3
- **ODN specific Plugins**: Set of common DPUs provided for Core usable only when
  whole ODN is used (for example DPUs used to load data and metadata specifically
to ODN/InternalCatalog)
  - https://github.com/OpenDataNode?query=UVPlugin
  - license: LGPLv3
- **Plugin Development Environment**: Tools and libraries needed by programmers for
development of new or modification of existing DPUs.
  - https://github.com/UnifiedViews/Plugin-DevEnv
  - license: combination of GPLv3 and LGPLv3
  - development of 3rd party DPUs with commercial licensing (i.e. non Open
    Source) possible thanks to LGPL
- **Packages**: Sources needed to build UnifiedViews packages for various Linux
distributions.
  - https://github.com/UnifiedViews/Packages
  - license: combination of GPLv3, LGPLv3 and Public Domain
Figure 4: Screenshot of UnifiedViews administration GUI showing list of data processing pipelines

Figure 5: Screenshot of UnifiedViews showing detailed configuration of one particular data processing pipeline, which consists of several DPUs
CKAN

COMSODE is reusing current implementation of CKAN as available in GitHub: https://github.com/ckan/ckan. For release management purposes, we’ve cloned the repository into our own copy: https://github.com/OpenDataNode/odn-ckan.

Licensing for CKAN remains as AGPLv3.0 and same applies for all changes and extensions COMSODE publishes.

All modifications and customizations required for ODN (i.e. for CKAN to serve as either public catalog or internal catalog: added functionality, updated look and feel, etc.) are done mostly as extensions, for easier maintenance and the further reuse. Modifications are kept at minimum and consist mostly of bug fixes (which we are not able to implement as extensions and which we are pushing for merge to upstream1).

Source code for both CKAN and our extensions can be found at GitHub under https://github.com/OpenDataNode/:

- https://github.com/OpenDataNode/odn-ckan - copy of official CKAN repository with small modifications
- https://github.com/OpenDataNode/kanext-odn-theme - ODN theme for CKAN
- https://github.com/OpenDataNode/kanext-odn-pipeline - extension implementing ODN pipeline association to dataset functionality
- https://github.com/OpenDataNode/kanext-odn-ic2pc-sync - extension for synchronization of catalog records from ODN/InternalCatalog to ODN/PublicCatalog, or to 3rd party (external) catalogs
- https://github.com/OpenDataNode/kanext-odn-cas - CAS authentication for CKAN
- https://github.com/OpenDataNode/odn-ckancommons - common software module needed for CKAN inside ODN

---

1 see for example https://github.com/ckan/ckan/pull/2392 and https://github.com/ckan/ckan/pull/2327
Figure 6: Dataset list in ODN/InternalCatalog, where publisher can manage all his datasets, both private and public
Figure 7: Screenshot showing integration of pipeline management between ODN/UnifiedViews and ODN/InternalCatalog: publisher can have basic overview and do basic task regarding dataset resources right from internal catalog.

**Single Sign-On**

ODN consists of several modules which are usable stand alone, are internal (i.e. accessible only to data publishers, not general public) and thus require authentication and authorization: ODN/UnifiedViews, ODN/InternalCatalog. When used in integrated product like ODN, it is very important for users to log into whole ODN just once: each user needs just one account for all ODN modules and login and logout needs to work from any module and be recognized by other modules seamlessly.

To achieve that, we implemented so called Single Sign-On in ODN and used following existing Open Source components:

- **midPoint**:
  - home page: [https://www.evolveum.com/midpoint/](https://www.evolveum.com/midpoint/)
  - license: APLv2
  - source code: [https://wiki.evolveum.com/display/midPoint/Source+Code](https://wiki.evolveum.com/display/midPoint/Source+Code)

- **CAS**:
  - home page: [https://www.apereo.org/projects/cas](https://www.apereo.org/projects/cas)
  - license: APLv2
  - source code: [https://github.com/Jasig/cas](https://github.com/Jasig/cas)

User database itself is stored in LDAP. midPoint is used to manage information about users stored in LDAP, allowing creation, modification and deletion of user accounts and other
additional features. CAS is sort of “integration glue”, which allows login and logout operations to work across several modules and use one central user database in LDAP.

Further changes to ODN/UnifiedViews and ODN/InternalCatalog (CKAN) were performed to allow those components to “plug” into midPoint + CAS infrastructure:

- in UnifiedViews, that’s part of the main source code now
- for CKAN we implemented it as extension: https://github.com/OpenDataNode/ckanext-odn-cas (mentioned also above in “CKAN” section)
  - Note: Given request for CAS authentication in official CKAN\(^2\), we plan to contribute or at least align our solution with that, which will be later developed for official CKAN.

This has also further benefits than just having Single Sign-On for whole ODN instance: IT administrators in bigger organizations can also take further advantage of this feature and integrate ODN with existing authentication and authorization services (for example via LDAP, Kerberos, Active Directory, SAML, etc.) so that users have same login and password for all IT services within organization (including ODN). This was utilized for example in eDemokracia pilot.

\(^2\) https://trello.com/c/2FlsJkFR/126-add-support-for-cas-authentication

![Login Screen](https://demo.comsode.eu/cas/login?service=https://bi.demo.comsode.eu)

**Figure 8:** Login screen for all internal parts of ODN. Internal parts (like internal catalog or UnifiedViews) are accessible only to publishers.
Figure 9: User management screen, where ODN administrator can manage users which can access internal parts of ODN

Packaging
COMSODE currently supports Debian Wheezy (7.x) as the main platform, where packaging and testing is performed. This decision is mainly motivated by CKAN primarily supporting that same platform.

“Source code” for packaging is available at GitHub too:
- [https://github.com/OpenDataNode/odn-box](https://github.com/OpenDataNode/odn-box) - for typical “single node” ODN installation
- [https://github.com/UnifiedViews/Packages](https://github.com/UnifiedViews/Packages) - for “stand alone” UnifiedViews (but reused in ODN)
- [https://github.com/OpenDataNode/odn-midpoint](https://github.com/OpenDataNode/odn-midpoint) - for midPoint component
- [https://github.com/OpenDataNode/odn-cas](https://github.com/OpenDataNode/odn-cas) - for CAS component

Licensing for packaging is same as licensing for each individual software module, i.e. Open Source in general.

EEA, as part of ODN release process, is building packages for each ODN release and making them available at:
[http://packages.comsode.eu/](http://packages.comsode.eu/)
Thanks to that, installation is easy:

1. Add ODN packages repository into apt-sources-list:
   ```bash
   echo "deb http://packages.comsode.eu/debian/ wheezy main" > /etc/apt/sources.list.d/odn.list
   ```

2. Add ODN public key:
   ```bash
   wget -O - http://packages.comsode.eu/key/odn.gpg.key | apt-key add -
   ```

3. Update apt sources:
   ```bash
   aptitude update
   ```

4. install ODN box:
   ```bash
   aptitude install odn-simple
   ```

(Please refer to https://github.com/OpenDataNode/open-data-node#installation for most up-to-date information.)

And upgrade even easier:

1. aptitude upgrade

Note: Additional operating systems are both planned and requested by potential users. See section "Future roadmap" on page 22.

**Virtuoso (Open Source version)**

COMSODE project reuses Open Source version of Virtuoso as is, without any modifications or customizations.

Virtuoso versions 7.2.x are supported by Open Data Node/UnifiedViews.


COMSODE provides our own (and according to some, even better\(^3\)) build of Virtuoso Open Source for Debian at http://packages.comsode.eu/pool/main/v/virtuoso-opensource/ which can be installed in a similar way as ODN, with only last command different:

```bash
apt-get install -y virtuoso-opensource=7.2
```

---

\(^3\) Official builds do not include isql tool. COMSODE build do include it. See http://serverfault.com/questions/631673/virtuoso-opensource-7-1-how-do-i-build-an-ubuntu-deb-package-that-includes-isql
5. Support infrastructure created for Open Data Node components

Support infrastructure for Open Data Node itself
COMSODE prepared following infrastructure for Open Data Node:
- public product web site: http://opendatanode.org/
- source code repository: https://github.com/OpenDataNode/
  - used also for tracking of issues and release management
- public Wiki: https://utopia.sk/wiki/display/ODN/Open+Data+Node+Home
  - used also for User and Administration Manuals: https://utopia.sk/wiki/display/ODN/Documentation
- public demo web site allowing visitors to “play” with ODN: http://demo.opendatanode.org/
- basic introductory presentation: http://www.slideshare.net/comsode/201504-odnplatformandmethodology

Support infrastructure for UnifiedViews
COMSODE contributed to preparation of following infrastructure for UnifiedViews:
- public product web site: http://www.unifiedviews.eu/
- source code repository: https://github.com/UnifiedViews
  - used also for tracking of issues and release management
- public Wiki: https://grips.semantic-web.at/display/UDDOC/Introduction
  - Administrator and Developer Guide: https://grips.semantic-web.at/display/UDDOC/UnifiedViews+Administrator+and+Developer+Guide
- mailing lists:
  - public: for broader community support: unifiedviews@googlegroups.com
  - private: for development collaboration: unifiedviews-dev@googlegroups.com
6. Services of Pilot publishing project

As part of Task 6.3 (Pilot publishing project), latest stable release of Open Data Node is deployed on COMSODE infrastructure and used by the consortium to publish 150⁴ COMSODE datasets.

For those datasets, instance of Open Data Node provides following publicly available services:

1. data catalog - available at http://data.comsode.eu/ listing all datasets published by consortium
2. data dumps for datasets - link available as resource for each such dataset in data catalog
3. REST API for “relational” datasets (data extracted from tabular sources like SQL databases and CSV or XLS files) - link available as resource for each such dataset in data catalog
4. SPARQL endpoint for “Linked Data” datasets - link available as resource for each such dataset in data catalog (generally at http://data.comsode.eu/sparql/)

Internally, COMSODE consortium utilizes also ETL and internal cataloguing functionality of this instance to publish datasets in a way similar to public bodies or some data re-users.

Data published this way is also used in search applications delivered by consortium and described in Deliverable 4.4.

This service (along with search applications), were providing real-world feedback for ODN development effort. EEA and Spinque plan to maintain the services for at least one more year, to support further showcase and Exploitation of COMSODE results.

Note: This section partially overlaps with Deliverable 6.4 (due next month), which will describe new datasets published by COMSODE project in more detail, including also some details about pilot instance of ODN.

---

⁴ Project is releasing new datasets continually, so the number of datasets reported on this catalog is increasing over time and will reach 150 or more by M22 (Jul 2015, when Deliverable D6.4 is due).
7. List of ODN releases

Thus the actual ODN releases delivered by COMSODE were as follows (main, feature releases are highlighted with bold):

<table>
<thead>
<tr>
<th>release</th>
<th>date</th>
<th>notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>test release D4.1</td>
<td>Oct 3rd 2014</td>
<td>only UnifiedViews 1.3.1 with some other modules</td>
</tr>
<tr>
<td>ODN 0.9.0</td>
<td>Dec 18th 2014</td>
<td>almost whole ODN, i.e. UnifiedViews and CKAN included, management module mainly missing</td>
</tr>
<tr>
<td>ODN 0.10.0</td>
<td>Feb 10th 2015</td>
<td>tighter integration of modules, management module still incomplete</td>
</tr>
<tr>
<td>ODN 0.10.1</td>
<td>Feb 11th, 2015</td>
<td>installation issue fixed</td>
</tr>
<tr>
<td>ODN 0.11.0</td>
<td>Mar 10th 2015</td>
<td>added synchronization with external data catalogs</td>
</tr>
<tr>
<td>ODN 1.0.1</td>
<td>Apr 29th 2015</td>
<td>whole ODN, including management module (user management and SSO), 1st stable release</td>
</tr>
<tr>
<td>ODN 1.0.2</td>
<td>May 29th 2015</td>
<td>bug-fix release</td>
</tr>
<tr>
<td>ODN 1.0.3</td>
<td>Jun 17th 2015</td>
<td>bug-fix release</td>
</tr>
</tbody>
</table>

For more up-to-date information please refer to [https://utopia.sk/wiki/display/ODN/Roadmap+and+releases](https://utopia.sk/wiki/display/ODN/Roadmap+and+releases).

At the end of M21 we can conclude that majority of planned functionality was implemented, although not everything went according to original plan (as described in section “Step 1 - development plan” in both this Deliverable and Deliverable 4.1):

- Some steps overlapped with others more, given relations between functions and use-cases, etc.
- Some steps took longer than anticipated, mainly because some risks actually materialized and countermeasures were not able to fully compensate (mainly staffing difficulties, see Periodic Progress Report for Year 1).
8. Future roadmap
As of M21, WP4 is officially over and ODN 1.0.3 (described here) is considered final. But further development work is still underway because:

- COMSODE is still launching some pilots (and other exploitation targets), which require agile response (regarding additional features) and support (bug fixes, etc.).
- ODN was from the beginning intended as market-ready Open Source platform. It is living software solution and development will continue in agile way after the end of the project because we believe that some of the pilots will transform into the business cases.
- WP4 accumulated some delays during Y1 which we were not able to fully counter during Y2, thus some features are still not implemented yet.

This continued development will be partially reported in Deliverable 6.3 and final project review (what was achieved, in which release, etc.).

This chapter gives overview of features which we plan still plan to implement within next few months:

1. as part of COMSODE (until M24, i.e. Sep 2015):
   a. integration of LDVMi into ODN
   b. commenting and other forms of feedback regarding published datasets
   c. cataloguing of applications (and their relations to datasets)
   d. further improvements to usability, including unification of look and feel in all modules
   e. further improvements to documentation, including proof-of-concept applications demonstrating use of data published via ODN via real runnable code
   f. ability to run scripts, scrapers, etc. within ETL pipelines, to accommodate existing or legacy publication solutions and non-Java tools

2. as part of further exploitation and business cases:
   a. DCAT support
   b. dataset versioning and publication via Git\(^5\), including integration with GitHub
   c. dataset usage statistics
   d. ability to harvest data from SOAP and REST APIs
   e. efficient replication of datasets using rsync\(^6\) protocol
   f. support for more platforms: newer Debian release (8.x), other Linux distributions (CentOS, …), OS X, Windows OS, bootable “live DVD”, etc.
   g. generic CAS based SSO solution for CKAN (see section “Single Sign-On”)
   h. and others, based on further feedback from pilots and potential clients

---

\(^5\) [http://www.git-scm.com/](http://www.git-scm.com/)