**H2020-INSO-2014**  
INSO-1-2014 ICT-Enabled open government  
YDS [645886] “Your Data Stories”

---

# D4.1: Customisation Plan v1.0

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Deliverable</td>
<td>D4.1: Customisation Plan 1.0</td>
</tr>
<tr>
<td>Workpackage</td>
<td>WP4: Customisation of Platform Modules</td>
</tr>
<tr>
<td>Nature</td>
<td>R</td>
</tr>
<tr>
<td>Dissemination Level</td>
<td>PU</td>
</tr>
<tr>
<td>Date</td>
<td>04/09/2015</td>
</tr>
<tr>
<td>Status</td>
<td>Final v1.0</td>
</tr>
<tr>
<td>Editor(s)</td>
<td>Georgios Petasis (NCSR-D)</td>
</tr>
<tr>
<td>Contributor(s)</td>
<td>L. Kallipolitis, A. Triantafillou (ATC)</td>
</tr>
<tr>
<td>Reviewer(s)</td>
<td>Paul Massey (TF), Bert Van Nuffelen (TF)</td>
</tr>
<tr>
<td>Document description</td>
<td>This report describes how existing and new technologies will be modified and optimised to support the YDS aims.</td>
</tr>
</tbody>
</table>
## Document Revision History

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Modifications Introduced</th>
<th>Modified by</th>
</tr>
</thead>
<tbody>
<tr>
<td>v0.1</td>
<td>25/05/2015</td>
<td>Initial Table of Contents</td>
<td>NCSR-D</td>
</tr>
<tr>
<td>v0.2</td>
<td>30/06/2015</td>
<td>Initial list of available solutions and infrastructures.</td>
<td>NCSR-D</td>
</tr>
<tr>
<td>v0.3</td>
<td>10/07/2015</td>
<td>Examination of capabilities of existing solutions.</td>
<td>NCSR-D</td>
</tr>
<tr>
<td>v0.4</td>
<td>24/07/2015</td>
<td>First draft of deliverable.</td>
<td>NCSR-D</td>
</tr>
<tr>
<td>v0.5</td>
<td>07/08/2015</td>
<td>Initial description of existing platforms and solutions.</td>
<td>NCSR-D</td>
</tr>
<tr>
<td>v0.6</td>
<td>14/08/2015</td>
<td>Updated images with higher resolution ones.</td>
<td>NCSR-D</td>
</tr>
<tr>
<td>v0.7</td>
<td>21/08/2015</td>
<td>Updated descriptions.</td>
<td>NCSR-D</td>
</tr>
<tr>
<td>v0.8</td>
<td>24/08/2015</td>
<td>Deliverable send for Quality Assurance.</td>
<td>NCSR-D</td>
</tr>
<tr>
<td>v0.9</td>
<td>03/09/2015</td>
<td>Quality Assurance review received.</td>
<td>TF</td>
</tr>
<tr>
<td>v1.0</td>
<td>04/09/2015</td>
<td>Final version &amp; Submission</td>
<td>NCSR-D, ATC</td>
</tr>
</tbody>
</table>
Executive Summary

This deliverable reports on the work performed in WP4/T4.1 regarding the preparation of the WP4 customisation planning, which addresses the design of the YourDataStories platform and applications. Having as a starting point the user requirements, as depicted in deliverable D2.3, “User Requirements v1.0”, and the YDS architecture as described in D2.5, “Technical Specifications and Architecture v1.0”, this deliverable presents a survey of the state-of-art regarding existing solutions and infrastructures that relate to the WP4 objectives. The user requirements from WP2 can be classified into three categories, relating into searching for data and within data, visualising data, and enhancing data, mainly through the addition of geo-location information. Based on these requirements, the survey focuses on existing solutions and infrastructures that provide search interfaces (such as data warehouses and initiatives that manage open data), visualise data, and provide geo-tagging functionality. The aim of the survey is to catalogue what functionalities are available, how information is searched, presented and visualised, and what are the technologies in use by the existing infrastructures and solutions available for open data. Finally, the deliverable includes the customisation plan/roadmap that will be followed for the development of WP4 components and applications.
Table of Contents

1 INTRODUCTION .................................................................................................................. 9
1.1 PURPOSE AND SCOPE ................................................................................................. 9
1.2 APPROACH FOR WORK PACKAGE AND RELATION TO OTHER WORK PACKAGES AND DELIVERABLES ................................................................. 9
1.3 METHODOLOGY AND STRUCTURE OF THE DELIVERABLE .................................. 10
2 APPLICATION TYPES ...................................................................................................... 11
3 STATE-OF-ART PLATFORMS ............................................................................................ 11
3.1 GEO-TAGGING ............................................................................................................. 11
3.1.1 Openstreetmap ...................................................................................................... 11
3.1.2 FixMyStreet .......................................................................................................... 13
3.1.3 Marine Traffic ....................................................................................................... 14
3.1.4 Cable map ............................................................................................................. 16
3.1.5 Ireland Road Collisions ....................................................................................... 17
3.2 DATA WAREHOUSES ................................................................................................ 18
3.2.1 Enigma ................................................................................................................. 18
3.2.2 Generation E ......................................................................................................... 20
3.2.3 Help me investigate ............................................................................................ 21
3.2.4 Public Data ........................................................................................................... 22
3.2.5 Quora ................................................................................................................... 23
3.2.6 Open Interests ..................................................................................................... 24
3.3 VISUALISATION/CHARTS ........................................................................................... 26
3.3.1 Openspending ..................................................................................................... 26
3.3.2 Public Spending ................................................................................................... 27
3.3.3 Statista ................................................................................................................ 29
3.3.4 World Bank ......................................................................................................... 30
3.3.5 Organization for Economic Co-operation and Development ............................ 31
3.3.6 Eurostat .............................................................................................................. 33
3.3.7 European Central Bank ..................................................................................... 36
3.3.8 Humanitarian Data Exchange .......................................................................... 36
3.3.9 The Atlas of Economic Complexity ................................................................ 38
3.3.10 Google Public Data Explorer ......................................................................... 43
3.3.11 SF OpenData .................................................................................................... 46
4 PROPOSED TECHNOLOGIES ........................................................................................... 49
4.1 YDS INFRASTRUCTURE .............................................................................................. 49
4.1.1 Content Management Systems .......................................................................... 49
4.1.1.1 Joomla ........................................................................................................ 49
4.1.1.2 Drupal ........................................................................................................ 49
4.1.1.3 Liferay ....................................................................................................... 49
4.1.2 Data Portals .......................................................................................................... 50
4.1.2.1 CKAN ....................................................................................................... 50
4.2 SEARCH FUNCTIONALITY ........................................................................................... 50
4.2.1 Databases ............................................................................................................. 50
4.2.1.1 MySQL ...................................................................................................... 50
4.2.1.2 MongoDB ................................................................................................. 50
4.2.2 API Infrastructure ................................................................................................ 50
4.2.2.1 Spring ....................................................................................................... 50
4.2.2.2 Laravel ..................................................................................................... 51
4.2.3 Search Engines ..................................................................................................... 51
4.2.3.1 Apache Solr ............................................................................................... 51
4.3 COMPONENTS ............................................................................................................. 51
4.3.1 User Interface ...................................................................................................... 51
## 4.3.1 jQuery / jQuery Mobile

- 4.3.1.1
- 4.3.1.2
- 4.3.1.3
- 4.3.1.4
- 4.3.1.5
- 4.3.1.6
- 4.3.1.7

## 4.3.2 AngularJS

- 4.3.2.1
- 4.3.2.2
- 4.3.2.3
- 4.3.2.4

## 4.3.3 Knockout.js

- 4.3.3.1
- 4.3.3.2
- 4.3.3.3
- 4.3.3.4
- 4.3.3.5

## 4.3.4 Ember.js

- 4.3.4.1
- 4.3.4.2
- 4.3.4.3
- 4.3.4.4
- 4.3.4.5

## 4.3.5 Bootstrap

- 4.3.5.1
- 4.3.5.2
- 4.3.5.3
- 4.3.5.4
- 4.3.5.5

## 4.3.6 Vaadin

- 4.3.6.1
- 4.3.6.2
- 4.3.6.3
- 4.3.6.4
- 4.3.6.5

## 4.3.7 Thymeleaf

- 4.3.7.1
- 4.3.7.2
- 4.3.7.3
- 4.3.7.4
- 4.3.7.5

## 4.3.8 Visualisations

- 4.3.8.1
- 4.3.8.2
- 4.3.8.3
- 4.3.8.4

## 4.3.9 Geo-tagging annotation facilities

- 4.3.9.1
- 4.3.9.2
- 4.3.9.3
- 4.3.9.4
- 4.3.9.5

---

## 5 Functionality Offered by the State-of-Art Approaches

## 6 Customisation Plan

## 7 Conclusions
List of Figures

Figure 1: Openstreetmaps's search page ................................................................. 12
Figure 2: Openstreetmap's page which shows GPS traces uploaded by users .................. 12
Figure 3: FixMyStreet's homepage ........................................................................ 13
Figure 4: FixMyStreet's search page ...................................................................... 14
Figure 5: Interactive map with vessel information (MarineTraffic) .............................. 15
Figure 6: Port information page (MarineTraffic) ....................................................... 15
Figure 7: Cable map's page which visualises the global undersea communication infrastructure .................................................. 16
Figure 8: Ireland Road Collisions' page which depicts collisions which took place in a specific area ................................................................. 17
Figure 9: Enigma's homepage ............................................................................... 18
Figure 10: Company information page (Enigma) ...................................................... 19
Figure 11: Generation E's homepage ..................................................................... 20
Figure 12: Search result page (HelpMeInvestigate) .................................................... 21
Figure 13: The search page of Public Data platform ................................................. 22
Figure 14: Graph visualisation chart (Public Data) ................................................... 23
Figure 15: Quora’s news feed ............................................................................. 24
Figure 16: OpenInterests' search screen ................................................................. 25
Figure 17: Openspending's screen for the selection of data visualisation type .............. 26
Figure 18: Openspending’s bubble tree visualisation chart ....................................... 27
Figure 19: PublicSpending’s homepage ................................................................. 28
Figure 20: The top expenditure categories of a country (PublicSpending) .................... 28
Figure 21: Infographic (Statista) .......................................................................... 29
Figure 22: Map visualisation (World Bank) ............................................................ 30
Figure 23: Line chart visualisation (World Bank) ..................................................... 31
Figure 24: Collection of visualisation indicators for Greece (OECD) ......................... 32
Figure 25: Global education spending indicator chart (OECD) .................................. 32
Figure 26: Economic Indicators (GDP per inhabitant) for Greece .............................. 33
Figure 27: Bubble chart showing total government revenue for Greece ...................... 34
Figure 28: Pie Charts showing quality of life for Greece. ......................................... 34
Figure 29: Android app displaying EU Economy statistics ........................................ 35
Figure 30: Eurostat Quiz (android app) .................................................................. 35
Figure 31: Eurostat Country profiles (android application) ........................................ 35
Figure 32: Line chart visualisation of ECB dataset .................................................. 36
Figure 33: Visualised datasets of a country (Humanitarian Data Exchange) ............... 37
Figure 34: Bar chart visualisation (Humanitarian Data Exchange) ............................ 37
Figure 35: Overview and line chart visualisation (Humanitarian Data Exchange) ...... 38
Figure 36: Tree map visualisation (Atlas of Economic Complexity) ........................... 39
Figure 37: Map visualisation (Atlas of Economic Complexity) .................................. 40
Figure 38: Stacked area visualisation (Atlas of Economic Complexity) ..................... 41
Figure 39: Force directed graph visualisation (Atlas of Economic Complexity) ......... 42
Figure 40: Bubble chart visualisation (Atlas of Economic Complexity) ..................... 43
Figure 41: Line chart visualisation (Google Public Data Explorer) ............................. 44
Figure 42: Bubble chart visualisation (Google Public Data Explorer) ....................... 44
Figure 43: Map visualisation (Google Public Data Explorer) .................................... 45
Figure 44: Bar chart visualisation (Google Public Data Explorer) ............................ 45
Figure 45: SF OpenData platform’s homepage ..................................................... 46
Figure 46: Table view (SF OpenData)................................................................. 47
Figure 47: Bubble chart visualisation (SF OpenData)................................................. 47
Figure 48: Map visualisation (SF OpenData)............................................................. 48
# List of Terms and Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>YDS</td>
<td>Your Data Stories</td>
</tr>
<tr>
<td>UI</td>
<td>User Interface</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>AIS</td>
<td>Automatic Identification System</td>
</tr>
<tr>
<td>KML</td>
<td>Keyhole Markup Language</td>
</tr>
<tr>
<td>REST</td>
<td>Representational State Transfer</td>
</tr>
<tr>
<td>API</td>
<td>Application Programming Interface</td>
</tr>
<tr>
<td>JSON</td>
<td>JavaScript Object Notation</td>
</tr>
<tr>
<td>SPARQL</td>
<td>Simple Protocol and RDF Query Language</td>
</tr>
<tr>
<td>PNG</td>
<td>Portable Network Graphics</td>
</tr>
<tr>
<td>XLS</td>
<td>Microsoft Excel spreadsheet file</td>
</tr>
<tr>
<td>PPT</td>
<td>Microsoft Powerpoint file</td>
</tr>
<tr>
<td>TXT</td>
<td>Text file</td>
</tr>
<tr>
<td>CSV</td>
<td>Comma Separated Values</td>
</tr>
<tr>
<td>PDF</td>
<td>Portable Document Format</td>
</tr>
<tr>
<td>DOC</td>
<td>Microsoft Word file</td>
</tr>
<tr>
<td>GEOJSON</td>
<td>Geographical JSON</td>
</tr>
<tr>
<td>KMZ</td>
<td>Compressed KML</td>
</tr>
<tr>
<td>BRC</td>
<td>British Red Cross</td>
</tr>
<tr>
<td>WFP</td>
<td>World Food Programme</td>
</tr>
<tr>
<td>SVG</td>
<td>Scalable Vector Graphics</td>
</tr>
<tr>
<td>CMS</td>
<td>Content Management System</td>
</tr>
<tr>
<td>PHP</td>
<td>PHP Hypertext Preprocessor</td>
</tr>
<tr>
<td>GNU</td>
<td>GNU’s Not UNIX</td>
</tr>
<tr>
<td>URL</td>
<td>Uniform Resource Locator</td>
</tr>
<tr>
<td>CKAN</td>
<td>Comprehensive Knowledge Archive Network</td>
</tr>
<tr>
<td>UNIX</td>
<td>Family of computer operating systems</td>
</tr>
<tr>
<td>BSON</td>
<td>Binary JSON format</td>
</tr>
<tr>
<td>POJO</td>
<td>Plain Old Java Object</td>
</tr>
<tr>
<td>SPA</td>
<td>Single Page Applications</td>
</tr>
<tr>
<td>MVVM</td>
<td>Model-View-View-Model</td>
</tr>
<tr>
<td>MVC</td>
<td>Model View Controller</td>
</tr>
<tr>
<td>HTML</td>
<td>HyperText Markup Language</td>
</tr>
<tr>
<td>DOM</td>
<td>Document Object Model</td>
</tr>
<tr>
<td>WebGL</td>
<td>Web Graphics Library</td>
</tr>
<tr>
<td>GDAL/OGR</td>
<td>Geospatial Data Abstraction Library</td>
</tr>
<tr>
<td>GML</td>
<td>Geographic Markup Language</td>
</tr>
<tr>
<td>TopoJSON</td>
<td>Extension of GeoJSON</td>
</tr>
<tr>
<td>SDW</td>
<td>Statistical Data Warehouse</td>
</tr>
<tr>
<td>ECB</td>
<td>European Central Bank</td>
</tr>
<tr>
<td>OECD</td>
<td>Organization for Economic Co-operation and Development</td>
</tr>
<tr>
<td>LESS</td>
<td>Dynamic style sheet language</td>
</tr>
</tbody>
</table>
1 Introduction

This is the first deliverable of task T4.1, “Customisation of Platform Modules”, regarding the preparation of the WP4 customisation planning, which addresses the design of the YourDataStories platform and applications. Having as a starting point the user requirements, as depicted in deliverable D2.3, “User Requirements v1.0”, and the YDS architecture as described in D2.5, “Technical Specifications and Architecture v1.0”, this deliverable presents a survey of the state-of-art regarding existing solutions and infrastructures that relate to the WP4 objectives, along with the customisation plan/roadmap that will be used for monitoring the development of the WP4 components and applications.

1.1 Purpose and Scope

This deliverable reports on the work performed in the context of task T4.1, “Customisation Planning”. The objectives of T4.1, which are also the objectives of this deliverable, are (as described in DoA):

1. Identify application types and categorise them into groups.
2. Identifying common components and subsystems to foster re-usability through the identification of functionalities which can be converted into re-usable building blocks.
3. Investigate the current state of the art, to identify mature or near to market technologies and standards, along with possible implementations that can be used and incorporated into the development of both the platform and the applications.
4. Define and maintain a plan for application development.

Having as a starting point the user requirements, as depicted in deliverable D2.3, “User Requirements v1.0”, and the YDS architecture as described in D2.5, “Technical Specifications and Architecture v1.0”, three categories have been identified regarding the functionality that must be supported by the components and applications of WP4 (objective 1), which are presented in section 2. These categories are used to identify, classify, and describe relevant projects, initiatives, infrastructures and applications, addressing objectives 2 and 3, through a survey that identifies commonalities in functionality and employed technologies by the existing applications surveyed (sections 3, 4, and 5). Finally, objective 4 is addressed in section 6.

1.2 Approach for Work Package and Relation to other Work Packages and Deliverables

The main responsibility of WP4 is to develop and deliver components readily usable in the YDS software stack, filling in all the levels of the YDS architecture beyond the Data Layer, which is the primary concern of WP3. WP4 is responsible for developing the YDS platform, all the required components and the pilot applications. It is organized in three main tasks, with T4.1 being responsible for examining the state-of-art in order to identify technologies and components that can be re-used or exploited within YDS. T4.2 is responsible for delivering the required components, such as components for personalisation, visualisation, the market-place, the help desk, and access control and identity management. Finally, task T4.3 is responsible for the YDS pilot applications provided through multiple interfaces and devices, such as mobile applications, web applications and applications related to social media.
Under this context, WP4 is of major importance with respect to the implementation of the YDS solution, as it implements the entire software stack that does not relate to the management of data (supported by WP3). As such, WP4 receives input from almost all work packages: From WP2 the user requirements, usage scenarios, and technical specifications and architecture. From WP3 the data published through the public interfaces and services WP3 will offer to the rest of the YDS infrastructure, and from WP6 feedback from pilots deployment and evaluation. Finally, WP4 provides the platform, components, and applications to WP5, concerned with their integration, technical assessment and testing.

1.3 Methodology and Structure of the Deliverable

The work presented in this deliverable has as a starting point the user requirements, as depicted in deliverable D2.3, “User Requirements v1.0”, and the YDS architecture as described in D2.5, “Technical Specifications and Architecture v1.0”. These requirements define what WP4 must provide, and can be grouped into functionality related to search, visualisation, and enhancing data through annotation, mainly related to geo-spatial information. Based on these three application types (search, visualise, enhance), we investigated the current state of art, conducting a survey of existing projects, initiatives, infrastructures and applications that provide functionality related to any of the three application types. Existing projects, initiatives, infrastructures and applications were examined with respect of the functionality they provide, but also with respect to the technologies used in their implementation, collecting a set of technologies that can be potentially re-used in YDS. The proposed technologies, reported in section 4, contain a set of alternative technologies for each of the YDS platform aspects examined (user interface, visualisation, geo-tagging). Exploitation of these technologies is left to the rest of the WP4 tasks: aiming at reusing existing components and applications as much as possible, the technologies that will be exploited within T4.2 and T4.3 will be defined by the existing components the two tasks will decide to integrate in the YDS infrastructure.

The rest of the deliverable is organised in five main sections. Section 2 presents the application types that must be supported by WP4, both in components and applications, as derived from the user requirements, as identified by WP2. Section 3 presents existing projects, initiatives, infrastructures and applications, in the form of short descriptions and sample screenshots, capturing aspects of the offered functionality and visual appearance. Section 4, “Proposed Technologies”, contains the most promising and mature technologies existing solutions employ, organised in categories, as a proposed arsenal of technologies that can incorporated in YDS, for its platform. Section 5 presents a comparative analysis of existing solutions regarding the capabilities offered, while section 6 presents the customisation plan that will be used to orchestrate and monitor development in tasks T4.2 and T4.3. Finally, section 7 concludes this deliverable.
2 Application types

The main user requirements that were collected and analysed in WP2, D2.3, “User Requirements v1.0”, and concern WP4, can be classified into three categories:

- **Search**: Users should be able to browse the available data, under multiple facets offered by the data, should be able to search the YDS platform in order to locate sources and data through their metadata, and finally users should be able to search within sources, to locate specific data of interest. Under these requirements we examined the search functionalities offered by the variable existing projects, initiatives, infrastructures and applications, regarding the depth of search (metadata or data), the facets offered, the data browsing capabilities, and the employed visualisation and search facilities.

- **Visualise**: Users should be able to view desired data in suitable forms of visualisation (like tables and graphs). Visualisation of data may be possibly associated with available geo-spatial information, shown on maps. YDS must offer the ability to visualise and compare multiple aspects of data on a single table/graph, apply transformations and visualise their results, and finally have the ability to create infographics. Under these requirements we examined the visualisation components employed by the existing projects, initiatives, infrastructures and applications, their comparison functionality, their capabilities regarding transformations. In addition we examined platforms that offer visualisation of maps and geo-spatial information.

- **Enhance**: This final category represents that set of requirements related to enhancing the data with the help of users, through actions like commenting, evaluating sharing and update of the data. A large portion of data enhancement in YDS relates to annotating data with geo-spatial information, thus platforms for annotating information on maps are of particular interest for this set of user requirements.

3 State-of-Art Platforms

This section presents existing projects, initiatives, infrastructures and applications, their capabilities and functionalities, organised into the three main application types (described in section 2) of particular interest to YDS: geotagging (section 3.1), data management that provide browsing and search functionality (section 3.2), and visualisation (section 3.3).

3.1 Geo-tagging

3.1.1 Openstreetmap

Openstreetmap (www.openstreetmap.org) is a collaborative project which aims to create a free editable map of the world. It is an open mapping service built by a community of mappers that contribute and maintain data about roads, trails, cafés, railway stations, and much more all over the world. The user has the ability to pinpoint locations around the world and explore their geopolitical information, to use the provided routing/navigation service in order to find their way to the desired destination, to export a specific portion of the map, as well as to add layers on the map in order to collect information about the transport system or the cycling infrastructure of a specific area. Openstreetmap also encourages users to upload their GPS traces in order to enrich and refine the existing mapping service. Last but not least, all of its data are distributed under an open license which means that researchers and developers are able to reuse them for any purpose.
Public GPS traces

Figure 1: Openstreetmap's search page

Figure 2: Openstreetmap's page which shows GPS traces uploaded by users
3.1.2 FixMyStreet

FixMyStreet (www.fixmystreet.com) is a platform which encourages users to report local problems in the UK, even if they don't know whom those reports should go to. The user types a postcode or an area name, pinpoints the problem in the map in order to show exactly where the problem is, and finally provides a description, a category and a picture of the problem. Afterwards, FixMyStreet is responsible to deliver the report to the corresponding local council in order to take care of the problem. Apart from these, users have the ability to read all the available reports, leave updates or receive alerts for all the problems reported within a particular ward or council, or all the problems within a certain distance of a particular location.

Figure 3: FixMyStreet’s homepage.
3.1.3 Marine Traffic

MarineTraffic (www.marinetraffic.com) is a vessel-tracking service which depicts in real-time all the available vessels on a map. The users have the ability to acquire the geographical positions of various vessels around the world. Moreover, the users can collect information about their characteristics (flag, picture, length, breadth, ship type and draught) as well as their destination, their speed, the ports that have visited, and their position history, whenever possible. Marine Traffic also provides information about the port’s departures and arrival worldwide. Finally, it encourages users to contribute to their platform by sending data acquired from their own AIS-Receiving Stations or by uploading pictures.
Figure 5: Interactive map with vessel information (MarineTraffic)

Figure 6: Port information page (MarineTraffic).
3.1.4 Cable map
Cable map (cablemap.info) is an attempt to consolidate all the available information about the undersea communications infrastructure and visualise them on a map. The initial data was harvested from Wikipedia and further information was collected through web search. Apart from viewing the existing communication cables, the users have the ability to view the cables which will be installed in the future. They can also filter the depicted cables either by type of cable or by the time period which the cables went active and provide feedback for the refinement of the map. All the data used are available in raw and KML format.

Figure 7: Cable map's page which visualises the global undersea communication infrastructure.
3.1.5 Ireland Road Collisions
Ireland Road Collisions (rsa.ie/RSA/Road-Safety/Our-Research/Collision-Statistics/Ireland-Road-Collisions) provides statistics about road collisions that took place in Ireland and visualises them as pins on a map. More specifically, it provides information about the type of the collision and its level of severity. When the user selects a specific area of the country, a grid with the area collisions grouped by date and time appears, summarising the day and the time that the collisions occurred.

![Ireland road collisions](image)

Figure 8: Ireland Road Collisions' page which depicts collisions which took place in a specific area.
3.2 Data warehouses

3.2.1 Enigma

Enigma’s (enigma.io) public data platform unifies billions of data points from more than 5,000 local, regional and national level sources. It enables users to search across public datasets, discover new datasets that they did not know they existed, and explore the connections among them. Users, for example, can learn how many cars travelled through the Holland Tunnel during the last hour, what was in all those containers unloaded at the Port of Los Angeles this morning, and from whom our politicians are receiving money. The data are shown in tabular format and users can apply filters on the whole table, as well as on each cell or column. By selecting one of the table’s columns they can view some quick statistics, while by clicking on each cell they are able to perform a quick search based on the content of the cell. Users also have the ability to see rankings for each column which are being generated from a number of predefined metrics.

All the data hosted on Enigma, can be downloaded in raw format. Moreover, they can be accessed through their RESTful API, which allows users to download datasets, query metadata, or perform server side operations on tables in Enigma.

![Enigma's homepage](image)

*Figure 9: Enigma's homepage.*
Figure 10: Company information page (Enigma).
3.2.2 Generation E
Generation E (www.rbdata.gr/generatione) is a cross-border data journalism project. It aims at collecting stories from young people that are moving from the southern countries (Portugal, Spain, Italy, Greece) in order to stay in other countries of the world. They attempt to discover why do they leave, where do they mainly go, and what they do now.
Since September 2014, more than 2000 stories have been collected. A dataset with the results of their initial research was released by the end of November.
3.2.3 Help me investigate

Help Me Investigate (helpmeinvestigate.com) is a website that helps people and journalists who want to investigate topics of public interest. More specifically, a user can search investigations performed on 4 specific topics (Education, Health, Welfare, Olympics). In addition, the users can read tips which will make their research easier. Despite its popularity, in 2014 the website was put on ice in order to look at possible new collaborative investigation projects. Website’s data are still publicly available in various formats depending on each dataset.

Help Me Investigate

![Search result page](HelpMeInvestigate)

Figure 12: Search result page (HelpMeInvestigate).
3.2.4 Public Data

PublicData (publicdata.eu)\(^1\) is a Pan-European data portal, providing access to open, freely reusable datasets from local, regional and national public bodies across Europe. It overcomes language barriers by aggregating datasets from official websites of public bodies in different countries, in multiple languages. Moreover, it provides licensing information for datasets and metadata about datasets, so as data users know what they can and cannot do with the data. The available datasets can be filtered based on category, tag, country and file format. Uses have the ability to browse, explore and compare datasets by region, subject matter, and format, as well as to visualise their data using line, point, bar or column graphs. They also have access to a personalised view that enables them to follow datasets, groups and users relevant to their interests. Finally, PublicData.eu hosts a large collection of applications which exploit the available datasets and showcase the opportunities arisen from their reusability.

---

\(^1\) PublicData is a prototype and is expected to be replaced by the new Pan-European Open Dataportal (nov 2015).
3.2.5 Quora

Quora (www.quora.com) is a question-answering on-line platform, which aims to provide answers from “professionals” on various topics. During the registration on Quora, the users are asked to declare the categories of topics in which they are interested in, and optionally their expertise. After registration, the user is able to submit, up-vote and rate questions, as well as to answer questions posted by other users. The questions are in free text format, and are always answered by human “experts”, i.e. by people who have declared that they have experience on the field that the question belongs in.

The interface is a feed-like one, where the topics shown are related to the user’s interests and sorted according to their popularity (up-votes, comments, etc.). Finally, the platform offers a twitter-like “trending now” feature, where the most trending groups of topics are shown.
OpenInterests (openinterests.eu) is a catalogue of political and commercial actors related to the European Union. It combines different sets of information into a search engine, which can be used to quickly retrieve information about the activities of companies, people and institutions in a European context. Built with Grano (granoproject.org), an open source tool for journalists and researchers who want to track networks of political or economic interest, it helps users to understand the most relevant relationships in their investigations, and to merge data from different sources. In addition, the users have the ability to export the retrieved information in JSON format.

Figure 15: Quora’s news feed.
European Parliament, Official Mail Service

Contracts awarded to companies


- Contract award 2012/S 77-125960 with a contract value of £2,216,080 for provision of specialised services: 1) Inspection firm; 2) ‘fire safety system’ coordinator (‘fire safety system’ coordinator) to Omnium Technique Européen (OTE)

- Contract award 2012/S 77-125960 with a contract value of £2,384,820 for provision of specialised services: 1) inspection firm; 2) ‘fire safety system’ coordinator (inspection firm) to Bureau Veritas

- Contract award 2012/S 193-316727 with a contract value of 655,384 for strengthening external perimeter protection — European Parliament, Louise Weiss Building (roads and other networks (No 01/100)) to Roessel SAS

- Contract award 2012/S 193-316727 with a contract value of £333,360.54 for strengthening external perimeter protection — European Parliament, Louise Weiss Building (natural stone work (No 04/102C)) to Demathieu & Bard

- Contract award 2012/S 193-316727 with a contract value of £277,861.20 for strengthening external perimeter protection — European Parliament, Louise Weiss Building (metalwork — anti-riot fencing (No 05/105)) to Heda SAS

- Contract award 2012/S 193-316727 with a contract value of £159,845.22 for strengthening external perimeter protection — European Parliament, Louise Weiss Building (carcassing (No 02/13D)) to CBA

Figure 16: OpenInterests' search screen.
3.3 Visualisation/Charts

3.3.1 Openspending

Openspending (openspending.org) is an open platform which aims to track and analyse public financial information globally. In more detail, it is a central, high-quality, open database of public financial information, including budgets, spending, balance sheets, procurement etc. All the available datasets include various information and metadata (source, timestamp, format, unit of measurement, and a team which is responsible for their data consistency). They also provide details about the different aspects that the datasets covered, such as the spender, recipient or classifications that may apply. The users are able to explore the available datasets, apply filters on them and visualise data using “bubble trees” or “tree map” representations. The produced visualisations can be saved, exported or embedded on other websites. Finally, all project data is made available under the Open Database License or the Open Data Commons license.

Figure 17: Openspending’s screen for the selection of data visualisation type.
3.3.2 Public Spending

Publicspending (publicspending.net) focuses on engineering meaningful interconnections among public spending data. More precisely, it is an effort to demonstrate the usefulness of economic Linked Open Data. The data are being collected from various sources like governmental data portals and Dbpedia. Afterwards, they are being processed, analysed, corrected if necessary, interlinked and depicted. The users have the ability to view how the money of each government is being spent worldwide. In addition, they can search for payments that companies have received from the indexed governments. Publicspending also provides a chart which depicts the top expenditure categories per country. All the available data can be accessed through a SPARQL endpoint, but if users want to
perform more complex queries or export the data, they have to register first, through the provided form.

Figure 19: PublicSpending’s homepage

Figure 20: The top expenditure categories of a country (PublicSpending).
### 3.3.3 Statista

Statista (www.statista.com) is a statistic portal which integrates data for over 80,000 topics from over 18,000 sources onto a single professional platform. Categorized into 21 market sectors, Statista provides companies, business customers, research institutions, and the academic community with direct access to quantitative data on media, business, finance, politics, and a wide variety of other areas of interest or markets. The data can be exported in various formats (PNG, XLS, PPT) or displayed in bar charts, line graphs or tabular formats. In addition, Statista offers a commercial service, which produces embeddable infographics on corporate-specific topics. Apart from that, the platform may also conduct studies, if requested, on an industry, or on a specific field which are available for download under a commercial license.

![Infographic: Which Country Sends The Most Remittances?](https://d28wbuch0jlv7v.cloudfront.net/images/infografik/normal/chartoftheday_3737 WHICH_COUNTRY_SEND_S_150701.png)

**Figure 21:** infographic (Statista).

<table>
<thead>
<tr>
<th>Country</th>
<th>Remittances (billion U.S. dollars)</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>$131 bn</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>$45 bn</td>
</tr>
<tr>
<td>United Arab Emirates</td>
<td>$29 bn</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>$25 bn</td>
</tr>
<tr>
<td>Germany</td>
<td>$24 bn</td>
</tr>
<tr>
<td>Canada</td>
<td>$23 bn</td>
</tr>
<tr>
<td>France</td>
<td>$21 bn</td>
</tr>
<tr>
<td>Russia</td>
<td>$21 bn</td>
</tr>
<tr>
<td>Italy</td>
<td>$16 bn</td>
</tr>
<tr>
<td>Spain</td>
<td>$16 bn</td>
</tr>
</tbody>
</table>

HTML code to embed chart [FAQ](http://www.statista.com/chart/3737/which-country-sends-the-most-remittances/)

```html
<a href="http://www.statista.com/chart/3737/which-country-sends-the-most-remittances/"
```

---

**Share on Facebook**

**Share on Twitter**

**Share on LinkedIn**

**Share on Google+**
3.3.4 World Bank

World Bank (www.worldbank.org) is a platform which provides free and open access to data about development in countries around the globe. The available datasets can be categorised by country or by topic (agriculture, health, poverty, etc.). Some of datasets are corrected using statistical techniques in order to allow users to perform meaningful comparisons across countries (indicators). In addition, World Bank offers a micro-data library which facilitates access to data collected through sample surveys of households, business establishments or other facilities. Depending on each dataset, users have the ability to apply custom filters and select from 3 different types of visualisation (graph, map, or table). The users can also embed the produced visualisations on their websites. Finally, country, topic and indicator pages allow the option to download all displayed data in common data exchange formats such as XLS, CSV and XML.

Life expectancy at birth, total (years)

Figure 22: Map visualisation (World Bank).
Life expectancy at birth, total (years)

Life expectancy at birth indicates the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life.


License: Open
Catalog: Sources World Development Indicators

![Line chart visualisation](World Bank).

3.3.5 Organization for Economic Co-operation and Development

OECD (www.oecd.org) is an organization that provides a forum in which governments can work together to share experiences and seek solutions to common problems. The organization works with governments to understand what drives economic, social and environmental change. It analyses and compares data to predict future trends, and therefore sets international standards on a wide range of things, from agriculture and tax to the safety of chemicals.

The available datasets can be categorised by country or by topic (agriculture, health, poverty, etc.). Depending on each dataset, users have the ability to apply custom filters and select from three different types of visualisation (graph, map, or table). The users can also embed the produced visualisations on their websites. Finally, country, topic and indicator pages can be freely downloaded in CSV format.
Figure 24: Collection of visualisation indicators for Greece (OECD).

Education resources

These indicators present data on education resources, including teachers’ and principals’ age, experience and the percentage of female staff, the number of students and staff teaching at different levels of education, staff hours spent in lesson preparation and teaching at lower secondary level, and teachers’ salaries. Data are also available for public and private spending. Public spending in...

Figure 25: Global education spending indicator chart (OECD)
3.3.6 Eurostat

Eurostat is the statistical office of the European Union situated in Luxembourg. Its task is to provide the European Union with statistics at European level that enable comparisons between countries and regions. Eurostat has implemented several data visualisation and mobile apps such as:

**Visualisation**

- The interactive data visualisation tool "Economic trends" provides every user, from basic to professional, a quick and easy overview of the development of 8 main economic indicators.
- The interactive data visualisation tool "My country in a bubble" allows you to compare your country with other European countries for more than 50 statistical indicators.
- "Quality of life" is a tool which provides you an insight into the different aspects of quality of life in your country and gives you the possibility to compare yourself with fellow citizens in your country as well as to the EU average and other EU Member States.

**Mobile apps**

- The EU Economy app gives mobile access to the most important short-term macroeconomic indicators (Principal European Economic Indicators - PEEIs) for the euro area, the EU and its Member States.
- Eurostat Quiz is a fun app to test your and your friends' knowledge about the EU and its Member States.
- The Eurostat Country Profiles application gives mobile access to the main statistical data for the EU and its Member States, the euro area, EFTA countries, EU candidate countries and a few other countries.

![Economic Indicators (GDP per inhabitant) for Greece.](image)

**Figure 26: Economic Indicators (GDP per inhabitant) for Greece.**
Figure 27: Bubble chart showing total government revenue for Greece.

Figure 28: Pie Charts showing quality of life for Greece.
Figure 29: Android app displaying EU Economy statistics.

Figure 30: Eurostat Quiz (android app).

Figure 31: Eurostat Country profiles (android application).
3.3.7 European Central Bank

ECB (www.ecb.europa.eu) is the central bank for the euro and administers monetary policy of the Eurozone. Along with other services, ECB offers Statistics Data Services that give users access to the euro area statistics, including in some cases national breakdowns. Users have the ability to search, display and save statistical series on-line using the “Statistical Data Warehouse”. More specifically, each dataset contains rich metadata (dataset explanation, contact information, quality references, etc.) in which users may apply highly customisable filters. Users are able to view the raw data in tables as well as to visualise them using line charts. Finally, they can subscribe to datasets in order to receive e-mail notifications when data are updated and export the available datasets in XML and CSV format.

![Statistical Data Warehouse](image)

Figure 32: Line chart visualisation of ECB dataset.

3.3.8 Humanitarian Data Exchange

Humanitarian Data Exchange (data.hdx.rwlabs.org) hosts a large collection of humanitarian data. The available datasets are divided in two categories, user-contributed datasets and curated indicators which have gone through a quality control process in order to be comparable across countries.
Users have the ability to explore the whole Humanitarian Data Exchange (HDX) collection by:

- Location (Greece, Italy, Spain, etc.)
- Topic (Economy, Education, Food and Nutrition, etc.)
- Organization (BRC Maps Team, WFP, etc.)

Additional filters (tags, license, format, date range) can be applied for a more fine-grained search. The format (TXT, CSV, XLS, PDF, JSON, DOC, GEOJSON, KML, KMZ, etc.) and the license of each dataset are defined by its owner. Although HDX’s datasets are mainly visualised using plots and bar charts, there are datasets which combine the aforementioned visualisation techniques with interactive maps, in order to give a more detailed point of view to the user.

![Map of Greece](image1)

**Figure 33: Visualised datasets of a country (Humanitarian Data Exchange).**

![Population density chart](image2)

**Figure 34: Bar chart visualisation (Humanitarian Data Exchange).**
3.3.9 The Atlas of Economic Complexity

The Atlas of Economic Complexity (atlas.cid.harvard.edu) is a powerful interactive tool that enables users to visualise a country’s total trade, track how these dynamically change over time, and explore growth opportunities for more than a hundred countries worldwide. More precisely, the user can discover what does a county import and export, how its trade evolves over time, what are the drivers of export growth and which industries are likely to emerge or disappear in a given territory. Atlas also provides estimates about the gross domestic product (GDP) growth rate of each country in the near future.

In order to search through the data, the users have to define the country in which they are interested, the desired time period, and the type of trade (import/export). More specific queries can be executed based on each product. The user has the ability to choose between many visualisation options. The trade data are being visualised using map, tree map and stacked visualisations, while there are more
complex visualisations available too. Finally, there is an option, which allows users to download each of the produced visualisations in PDF, PNG, and SVG formats.

Figure 36: Tree map visualisation (Atlas of Economic Complexity).
Figure 37: Map visualisation (Atlas of Economic Complexity).
Figure 38: Stacked area visualisation (Atlas of Economic Complexity).
Figure 39: Force directed graph visualisation (Atlas of Economic Complexity).
3.3.10 Google Public Data Explorer

Google Public Data Explorer (www.google.com/publicdata) holds a large collection of datasets and metrics acquired from various sources (Eurostat, World Bank, World Economic Forum, etc.). The datasets are being categorised based on the language published. Filtering can be applied, by specifying the data provider. Depending on the dataset, users have also the ability to compare the data included (i.e. per country, gender, marital status) by using the available data visualisations (line, bar, bubble, map charts and tree maps). In addition, the users can upload their own datasets and apply different types of visualisations. The produced visualisations can be shared as well as to be embedded in other websites.
Figure 41: Line chart visualisation (Google Public Data Explorer).

Figure 42: Bubble chart visualisation (Google Public Data Explorer).
Figure 43: Map visualisation (Google Public Data Explorer).

Figure 44: Bar chart visualisation (Google Public Data Explorer).
3.3.11 SF OpenData

SF OpenData (data.sfgov.org) is the City and County of San Francisco’s official open data portal. It contains hundreds of city datasets grouped into 10 main categories (Economy & Community, City Management & Ethics, Transportation, Public Safety, Health & Social Services, Geographic Locations & Boundaries, Energy & Environment, Housing & Buildings, City Infrastructure, Culture & Recreation) for use by developers, analysts, residents, and more. Depending on the dataset, the user can choose between several visualisation options (line, bar, pie, donut, line, area, timeline, bubble, or tree map charts). Filters can also be applied by the user in order to get the most out of each dataset. Finally, the user has the ability to comment on the dataset, export the dataset in various formats and embed the produced visualisations.

![Figure 45: SF OpenData platform’s homepage.](image-url)
Figure 46: Table view (SF OpenData).

Figure 47: Bubble chart visualisation (SF OpenData).
Figure 48: Map visualisation (SF OpenData).
4 Proposed Technologies

This section contains the set of mature and near-to-market technologies, which are used by the existing projects, initiatives and applications that have been described in section 3. They are organised around the three application types described in section 2 (search – section 4.2, visualise – section 4.3.2, and enhance – section 4.3.3), but also along aspects of the infrastructure like the platform (section 4.1) and the user interface (section 4.3.1). Exploitation of the technologies reported in the section is attributed to the rest of the WP4 tasks: aiming at reusing existing components and applications as much as possible, the technologies that will be exploited within T4.2 and T4.3 will be defined by the existing components the two tasks will decide to integrate in the YDS infrastructure.

4.1 YDS Infrastructure

4.1.1 Content Management Systems

4.1.1.1 Joomla

Joomla (joomla.org) is an open-source content management system (CMS) written in PHP, which enables users to build feature-rich websites and powerful online applications. It is capable of carrying out tasks ranging from corporate websites and blogs to social networks and e-commerce. Joomla is built on a model–view–controller (MVC) web application framework that can be used independently of the CMS. Since its initial release, Joomla’s primary focus has been on usability and extensibility. One of the key benefits of Joomla is the module customisation option. With highly customisable site modules (and extensions), Joomla can support almost any functionality. Moreover its advanced content management system enables users to organise their content in any way. Finally, Joomla features user management, media manager, contact manager as well as a template management system which allows users to easily change the look n feel of the website without distorting its original content.

4.1.1.2 Drupal

Drupal (drupal.org) is an open-source content-management framework written in PHP and distributed under the GNU open source license. It is used as a back-end framework for various types of websites ranging from personal blogs to corporate, political, and government sites. Drupal allows users to easily organise, manage and publish their content with an endless variety of customisation. Its content management system allows users to categorise their content through URL addresses, paths, making their own lists. This structure makes easier the management, search and reuse of the content. Moreover, Drupal’s presentation layer allows designers to create highly usable and interactive experiences. It supports collaborative authoring, newsletters, podcasts, image galleries, peer-to-peer networking, file uploads/downloads and more. The base functionality provided by the Drupal’s platform can be further extended by using one of the thousands available plugins or by developing a new one that will cover the needs of every project.

4.1.1.3 Liferay

Liferay (liferay.com) is a free and open source enterprise portal project written in Java and distributed under the GNU open source license and (optionally) under a commercial license. It is a web platform with features commonly required for the development of websites and portals. It includes a built-in web content management system, allowing users to build websites and portals as an assembly of themes, pages, portlets/gadgets, and a common navigation. Liferay’s support for plugins extends into
multiple programming languages, including support for PHP and Ruby portlets. Although it offers a sophisticated programming interface for developers, no programming skills are required for basic website installation and administration.

4.1.2 Data Portals

4.1.2.1 CKAN
CKAN (ckan.org) is a powerful data management system that makes data accessible by providing tools to streamline publishing, sharing, finding and using data. It is aimed at data publishers (national and regional governments, companies and organisations) wanting to make their data open and available. CKAN is built with Python on the backend and JavaScript on the frontend. It has a modular architecture that allows extensions to be developed to provide additional features such as harvesting or data upload. CKAN uses its internal model to store metadata about the different records, and presents it on a web interface that allows users to browse and search these metadata. It also offers a powerful API that allows third-party applications and services to be built around it.

4.2 Search Functionality

4.2.1 Databases

4.2.1.1 MySQL
MySQL (mysql.com) is an open source, full-featured relational database management system (RDBMS). It is based on the Structure Query Language (SQL) and is queried using a subset of the standard SQL commands. MySQL runs on many different platforms, including Linux, UNIX, and Windows. Although it can be used in a wide range of applications, MySQL is most often associated with web-based applications and online publishing. It is designed to be fully multi-threaded using kernel threads, to easily use multiple CPUs if they are available, and provides transactional and non-transactional storage engines. It is relatively easy to add other storage engines. Finally, MySQL uses a very fast thread-based memory allocation system which makes the queries that require combining more than one table to run faster and more efficiently.

4.2.1.2 MongoDB
MongoDB (mongodb.org) is a cross-platform, document-oriented, database that provides high performance, high availability, and easy scalability. Classified as a NoSQL database, MongoDB eschews the traditional table-based relational database structure in favour of JSON-like documents with dynamic schemas (MongoDB calls the format BSON), making the integration of data in certain types of applications easier and faster. Documents comprise sets of key-value pairs and are the basic units of data within MongoDB. Collections contain sets of documents and function as the equivalent of relational database tables. Its automatic sharding enables data in a collection to be distributed across multiple systems for horizontal scalability as data volumes increase.

4.2.2 API Infrastructure

4.2.2.1 Spring
The Spring Framework (projects.spring.io/spring-framework) is an open source Java platform that provides comprehensive infrastructure support for developing Java applications which run on any kind of deployment platform. The core features of the Spring Framework can be used in developing any Java application, but there are extensions for building web applications on top of the Java EE
platform. Spring framework targets in making J2EE development easier to use, and to promote good programming practices by enabling a POJO-based programming model. The Spring framework is a layered architecture, which consists of several modules. Its modular architecture enables developers to eliminate the modules which are of no use.

### 4.2.2.2 Laravel

Laravel (laravel.com) is an open source PHP web application framework intended for the development of web applications following the Model–View–Controller (MVC) architectural pattern. Laravel is accessible, yet powerful, providing powerful tools needed for large, robust applications. It attempts to simplify the development cycle by easing common tasks used in the majority of web projects, such as authentication, routing, sessions, and caching. In more detail, Laravel features expressive syntax, a modular packaging system with a dedicated dependency manager, different ways for accessing relational databases, and various utilities that aid in application deployment and maintenance.

### 4.2.3 Search Engines

#### 4.2.3.1 Apache Solr

Apache Solr (lucene.apache.org/solr) is an open source enterprise search platform, written in Java. It include full-text search, hit highlighting, faceted search, near real-time indexing, dynamic clustering, database integration, NoSQL features and rich document (e.g., Word, PDF) handling. Providing distributed search and index replication, Solr is designed for scalability and fault tolerance. It runs as a standalone full-text search server. Moreover, it uses the Lucene Java search library at its core for full-text indexing and search, and has REST-like HTTP/XML and JSON APIs that make it usable from most popular programming languages. Solr’s external configuration allows it to be tailored to many types of application without Java coding, and it has a plugin architecture to support more advanced customisation.

### 4.3 Components

#### 4.3.1 User Interface

##### 4.3.1.1 jQuery / jQuery Mobile

jQuery (jquery.com) is a fast, small, and feature-rich JavaScript library. It makes things like HTML document traversal and manipulation, event handling, animation, and Ajax much simpler with an easy-to-use API that works across a multitude of browsers. jQuery is a free, open source, and dual-licensed library under the GNU General Public License. It offers functionality that allows developers to build plug-ins, in addition to the JavaScript library. This allows for the development of abstractions for animation and low-level interaction, sophisticated effects and theme-able, high-level widgets. The modular mechanism of the jQuery library facilitates the development of highly effective, potent Web applications and Web pages. Finally, there is a mobile version of jQuery which is designed to make easier the development of responsive web sites and apps that are accessible on all smartphone, tablet and desktop devices.

##### 4.3.1.2 AngularJS

AngularJS (angularjs.org) is an open-source web application framework maintained by Google which attempts to address many of the challenges encountered in developing SPAs. It aims to simplify both the development and the testing of such applications by providing a framework for client-side MVC and MVVM architectures, along with components commonly used in rich Internet applications. The
framework adapts and extends traditional HTML to present dynamic content through two-way data-binding that allows for the automatic synchronisation of models and views. As a result, AngularJS de-emphasizes explicit DOM manipulation with the goal of improving testability and performance.

4.3.1.3 Knockout.js
Knockout (knockoutjs.com) is an open source JavaScript library which implements Model-View-View Model (MVVM) design pattern and helps the creation of rich, desktop-like web user interfaces (UIs). It simplifies user interactions and makes interfaces fully responsive to any data source changes. More specifically, it includes declarative bindings that allow developers to bind the elements of UI to the data model in a simple and convenient way, automatic UI refresh when the data model's state changes, dependency tracking and a native templating engine. In general, Knockout follows Angular’s philosophy but it lacks important features which are natively supported by Angular like routing/navigation support and services for working with network calls like REST services.

4.3.1.4 Ember.js
Ember.js (emberjs.com) is an open-source JavaScript application framework, based on the model-view-controller (MVC) pattern. It allows developers to create complex single-page web applications which are highly scalable, by incorporating common idioms and best practices into the framework. Although primarily considered a framework for the web, it is also possible to build desktop and mobile applications in Ember. It is well suited for applications that display dynamic data and have increased user interaction. Such applications include task managers, dashboards, forums, chat and messaging applications, and so on. Ember comes tightly integrated with a templating engine known as “Handlebars”, which gives Ember one of its most powerful features: two-way data-binding. It also offers other features such as state management, auto-updating templates and computed properties.

4.3.1.5 Bootstrap
Bootstrap (getbootstrap.com) is a front-end development framework that enables developers and designers to quickly build fully responsive websites. It is compatible with the latest versions of the Google Chrome, Firefox, Internet Explorer, Opera, and Safari browsers, although some of these browsers are not supported on all platforms. Bootstrap is modular and consists essentially of a series of LESS stylesheets that implement the various components of the toolkit. These include buttons with advanced features (e.g. grouping of buttons or buttons with drop-down option, make and navigation lists, horizontal and vertical tabs, navigation, breadcrumb navigation, pagination, etc.), labels, advanced typographic capabilities, thumbnails, warning messages and a progress bar. The components are implemented as CSS classes, which must be applied on certain HTML elements in a page. Developers are able to build custom configurations of the Bootstrap file in order to include only the desired components.

4.3.1.6 Vaadin
Vaadin (vaadin.com) is an open source web application framework built on Java. It is designed for creating rich and interactive applications that run in the browser, without any plugins. A server-driven architecture together with reusable component model is used to simplify programming of applications and for better web application security. The programming model is much the same as in traditional desktop programming, with events and listeners rather than requests and responses. Vaadin uses Google Web Toolkit for rendering the resulting web page. While the way Vaadin uses Google Web Toolkit could lead to trust issues, it only operates client-side and Vaadin adds server-side data validation to all actions. This means that if the client data is tampered with, the server notices this and doesn’t allow it.
4.3.1.7 Thymeleaf

Thymeleaf (thymeleaf.org) is a Java XML/XHTML/HTML5 template engine which is able to apply a set of transformations to template files in order to display data and/or text produced by the applications. It can work both in web (Servlet-based) and non-web environments. Thymeleaf is better suited for serving XHTML/HTML5 at the view layer of MVC-based web applications, but it can process any XML file even in offline environments. The main goal of Thymeleaf is to provide an elegant and well-formed way of creating templates. In order to achieve this, it is based on XML tags and attributes that define the execution of predefined logic on the DOM, instead of explicitly writing that logic as code inside the template. Its architecture allows a fast processing of templates, relying on intelligent caching of parsed files in order to use the least possible amount of I/O operations during execution. Last but not least, Thymeleaf has been designed from the beginning with XML and Web standards in mind, allowing developers to create fully validating templates if needed.

4.3.2 Visualisations

4.3.2.1 Chart.js

Chart.js (chartjs.org) is dependency free and super lightweight JavaScript library which uses the HTML5 canvas element in order to produce different types of interactive data visualisations. It offers six different types of visualisations (line, bar, radar, polar area, pie, and doughnut chart), which can be customised according to the needs of each project. More specifically, it allows users to load only the chart types they need, and on top of that to enable or disable additional features for each chart. In addition, all the available chart types are responsive, which means that the charts will resize automatically to fit in the browser’s window size, along with providing the perfect scale granularity for that size.

4.3.2.2 Highcharts

Highcharts (highcharts.com) is a JavaScript library, which is solely based on native browser technologies and offers a wide range of interactive charts. It supports line, spline, area, areaspline, column, bar, pie, scatter, angular gauges, arearange, areasplinerange, columnrange, and polar chart types, while many of these can be combined in one chart. Through the provided API the user can add, remove and modify series and points, or modify axes at any time after chart creation. This feature gives users the ability to create live charts that will be constantly updated with new values from the server. Moreover, Highcharts lets users assign a y axis for each series - or an x axis in case they want to compare data sets of different categories. Last but not least, all the produced charts can be either exported to PNG, JPG, PDF or SVG formats or printed directly from the web page.

4.3.2.3 D3.js

D3.js (d3js.org) is the most popular JavaScript library for manipulating documents based on data. More precisely, it allows users to bind arbitrary data to a Document Object Model (DOM), and then apply data-driven transformations on the document. It allows users to visualise sizeable portions of data by using HTML, SVG, and CSS. D3 supports area, line, bivariate area, multi-series line, stacked area, bar, stacked bar, normalized stacked bar, grouped bar, donut, and pie chart types. Apart from these, users are able to create box plots, calendar views, chord diagrams, dendrograms, streamgraphs, force-directed graphs, treemaps, scatterplots, and much more. The aforementioned visualisations can be customised or extended in order to produce animations or depict more information about the visualised data according to the interactions of the user.
4.3.2.4 Dc.js
Dc.js (dc.js.github.io/dc.js) is a JavaScript charting library with native crossfilter (square.github.io/crossfilter) support which allows highly efficient exploration on large multi-dimensional datasets, which are depicted in desktop browsers as well as on mobile devices. It leverages d3 (d3js.org) engine to render charts in CSS friendly browsers. Charts rendered using dc.js are naturally data driven and reactive, therefore providing instant feedback on user’s interaction. More specifically, dc.js supports base, colour, stackable, coordinate grid, pie, row, bar, line, composite, abstract bubble, bubble, bubble overlay, and geo choropleth charts. Except from charts, it offers a couple of widgets (data count and data table widgets) which are designed to list crossfilter focused datasets. The use of such widgets makes much more efficient the filtering of the data that correspond to a specific point of a chart.

4.3.3 Geo-tagging annotation facilities

4.3.3.1 OpenLayers
OpenLayers (openlayers.org) is high-performance, feature-packed JavaScript library which allows users to create their own interactive maps. It leverages WebGL, Canvas 2D and other HTML5 features in order to deliver smooth user-experience on all major browsers, desktop and mobile devices. OpenLayers is capable of rendering vector data from GeoJSON, TopoJSON, KML, GML, and a growing number of other formats. Users have the ability to apply layers and add custom elements like pins, tooltip, tiles and map controls, which can be styled further with CSS. Finally, it is possible to use OpenLayers along with the D3 visualisation library.

4.3.3.2 Leaflet
Leaflet (leafletjs.com) is an open source JavaScript library for mobile-friendly interactive maps. It works efficiently across all major desktop and mobile platforms and delivers a smooth user experience. In order to achieve such levels of performance it utilises CSS3 features and hardware acceleration on mobile devices. Regarding its features, it allows users to apply custom map layers and controls, to add markers with their own icons, as well as to annotate specific areas on the map. It also enables users to annotate specific paths on the map and draw custom shapes (polygons, rectangles and circles) in order to highlight specific areas. Furthermore, the elements that are being added on the map can be styled using CSS3. In case that the core features does not cover the user needs, custom plugins can be employed in order to achieve the desired functionality.

4.3.3.3 Polymaps
Polymaps (polymaps.org) is a JavaScript library for making dynamic, interactive maps in modern web browsers. It provides speedy display of multi-zoom datasets over maps, and supports a variety of visual presentations for tiled vector data, in addition to the usual cartography from OpenStreetMap, CloudMade, Bing, and other providers of image-based web maps. Polymaps uses SVG to display information, so it is easy to define the design of the rendered data by using CSS rules.

4.3.3.4 Kartograph
Kartograph (kartograph.org) is a simple and lightweight framework for creating beautiful, interactive vector maps without Google Maps or any other mapping service. The framework provides two separate libraries. The first one is a Python library, which is built on top of GDAL/OGR (www.gdal.org) and shapely (toblerity.org/shapely), and enables users to generate beautiful, Illustrator-friendly SVG maps as well as to render vector maps from shapefiles and PostGIS. The other provided library is a
JavaScript library built on Raphael (raphaeljs.com) and jQuery (jquery.com) for creating interactive maps based on Kartograph.py SVG maps which run across all major browsers.

4.3.3.5 Mapbox
Mapbox (mapbox.com) is an enterprise mapping platform for developers which makes easy the integration of location services into any mobile or online application. More specifically, it allows developers to choose from a wide range of base maps and customize them according to their needs. Mapbox also offers the ability to create custom Visualisations for the data which will be depicted on the map. Moreover, it provides driving, walking, and cycling routes, with turn-by-turn directions. Apart from the aforementioned services, it incorporates a geocoding service which enables developers to turn their coordinates into addresses or addresses into coordinates with a simple call to the Mapbox’s API.
5 Functionality offered by the state-of-art approaches

This section aims to summarise the most prominent functionality that is offered by the platforms described in section 3. Furthermore, it also intends to examine which of these features may be potentially included in the YDS platform, if such a need arises.

One of the most important facilities that such kinds of platform have to offer is search. Search is of great importance, as it enables users to find the data that they are looking for without knowing the exact details of the desired dataset. The platforms mentioned in section 3 provide various types of search. Most of them allow users to perform search upon the fundamental information of the datasets while others offer more advanced search features. In more detail, platforms like “Google Public Data Explorer” and “World Bank” provide users with the ability to filter the available datasets based on their parent categories (country, organization etc.). “Enigma”, “Public Data” and “Humanitarian Data Exchange” enable users to query the dataset’s metadata for a more fine-grained search while “SF OpenData” goes one step further and allow users to search inside the dataset itself. Such tasks can be accomplished by using the CKAN data portal, which provides extensive support in describing data sources with metadata and searching within these metadata. Alternatively, Apache Solr is able to provide similar functionality by customising its search parameters.

Apart from search, another feature frequently provided by state-of-art approaches is data comparison, usually through visualisation of different data or aspects of data in a single chart. When datasets use the same units of measurement, this task can be easily accomplished with the most of the JavaScript data visualisation libraries described in section 4. Platforms like “Google Public Data Explorer” and “Humanitarian Data Exchange” already offer this feature. Comparison becomes more complicated when the units of measurement differ between datasets. For example, “World Bank” offers users the ability to compare the debt of countries with different currencies. This task requires datasets which have been pre-processed in order to be made comparable between each other, a transformation that depends on the data and may require significant resources (in enhancement and validation of the data) in order to be accomplished.

A core feature that several platforms, such as Google Public Data Explorer, PublicSpending, ECB, and OECD miss, is the enhancement of the datasets after their publication. It is quite frequently a desirable feature that the data included in each dataset is accurate, consistent and in a form that can be reusable by other users. Although such properties can be addressed with the contribution of users (i.e. feedback, evaluation, etc.), most platforms doesn’t offer relevant functionality. Only a limited number of the surveyed platforms (Enigma and SF OpenData) allow users to send explicitly their feedback for each dataset, while only “SF OpenData” allows users to comment on these.

Last but not least, open data have to be easily shared in order to be accessible from a growing number of users. Most of the platforms give users the ability to download datasets in different formats as well as to access them though the available APIs. Apart from that, platforms like “Humanitarian Data Exchange” and “The Atlas of Economic Complexity” allow users to post sharable links of the datasets in social networks like Facebook and Twitter.
6 Customisation Plan

The customisation plan has been designed according to the project’s milestones, divided per task and taking into account the effort estimated for each task. The plan is presented below.

**M7 (Aug 15) – M11 (Dec 15): Implementing the first version of components and applications**
The aim of the first version of components and applications is to mainly support the first version of the pilot applications, through the development of the components and applications that will be integrated into the first integrated prototype by WP5.

- M7-M10: Service marketplace infrastructure version 1.0
- M9-M11: Visualisation infrastructure version 1.0
- M9-M11: Applications and components version 1.0
- M10-M11: Personalization infrastructure version 1.0
- M10-M11: Identity management system and Help Desk version 1.0
- M11: Test cases and technical evaluation

**M12 (Jan 16) -M23 (Dec 16): Implementing the second version of components and applications**

- M12-M14: Service marketplace infrastructure version 2.0
- M14-M17: Visualisation infrastructure version 2.0
- M16-M19: Applications and components version 2.0
- M18-M21: Personalization infrastructure version 2.0
- M21-M23: Identity management system and Help Desk version 2.0
- M23: Test cases and technical evaluation

**M24 (Jan 17) - M32 (Sept 17): Implementing the final integrated prototype and the final version of applications**

- M24-M26: Service marketplace infrastructure final version
- M26-M29: Visualisation infrastructure final version
- M28-M31: Applications and components final version
- M29-M31: Personalization infrastructure final version
- M31-M32: Identity management system and Help Desk final version
- M32: Test cases and technical evaluation
| Task | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 |
| Customisation of Platform Modules |   |   |   |   |   |   |   |   |   | MS1: Version 1.0 |   |   |   |   |   |   |   |   |   |   |   |   | MS1: Version 1.0 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Service marketplace infrastructure |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Visualization infrastructure |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Applications and components |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Personalization infrastructure |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Identity management system and Help Desk |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Test cases and technical evaluation |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

MS1: Version 1.0
MS2: Version 2.0
MS3: Final Version
7 Conclusions

This deliverable has presented the work that has been performed in the context of task T4.1, of work package WP4. Having as a starting point the user requirements, as depicted in deliverable D2.3, “User Requirements v1.0”, and the YDS architecture as described in D2.5, “Technical Specifications and Architecture v1.0”, three categories have been identified regarding the functionality that must be supported by the components and applications of WP4. These categories have been used to identify, classify, and describe relevant projects, initiatives, infrastructures and applications, through a survey that identifies commonalities in functionality and employed technologies by the existing applications surveyed. Finally, a customisation plan/roadmap has been designed, in order to monitor the development process of WP4 components and applications.

Cataloguing the state-of-art regarding the existing solutions will familiarise WP4 development to what is already available, in terms of user interfaces, data visualisation, search facilities, data annotation with geo-spatial and other information, interfaces for embedding visualisation on external applications, and offered functionality on data. Along with the collected information about existing solutions, a list of state-of-art, mature, and near to market technologies were collected and briefly described, in order to form a set of proposed technologies that can be exploited by tasks T4.2 and T4.3, which are responsible for developing the YDS platform components and applications respectively. Exploitation of these technologies is left to the rest of the WP4 tasks: aiming at reusing existing components and applications as much as possible, the technologies that will be exploited within T4.2 and T4.3 will be defined by the existing components the two tasks will decide to integrate in the YDS infrastructure.

This deliverable is the first version of the customisation plan, which has addressed all requirements of task T4.1, reported in section 1.1. User requirements from WP2 were used to identify three application types, which guided the survey on state-of-art approaches on functionality and employed technologies. Subsequent versions are expected to provide an updated customisation plan and required adaptations to potential changes in the requirements from WP2.