



## Raising Open and User-friendly Transparency- Enabling Technologies for Public Administrations



Project number 645860  
H2020-INSO-2014

### D4.2 Alpha version Of TET

(Draft, version 1.0, 29/01/2016)



WISE&MUNRO



**Document produced by**

Organisation: Insight, NUIG

Author / email:

Adegboyega Ojo, adegboyega.ojo@insight-centre.org

Edobor Osagie, edobor.osagie@insight-centre.org

Mohamamd Waqar, mohammad.waqar@insight-centre.org

Arkadiusz Stasiewicz, arkadiusz.stasiewicz@insight-centre.org

Samuel Adebayo, samuel.adebayo@insight-centre.org

Lukasz Porwol, lukasz.porwol@insight-centre.org

Subject: Alpha version of TET

Due date: 31 January 2016

**Dissemination level:** Public (PU)

**Reviewed and approved by**

Date	Name	Organization
2016-01-24	Dr Vangelis Banos Ilias Trochidis	Ortlio Ltd
2016-01-25	Vittorio Scarano	UNISA
2016-01-25	Elena Palmisano Paolo Boscolo	Comune di Prato

**Revision History**

Version	Date	Authors	Status	Description of Changes
0.1	08-01-2016	A. Ojo	Draft	A first organization
0.2	12-01-2016	E. Osagie	Draft	Section 1 and 6
0.3	18-01-2016	M. Waqar	Draft	Section 2 and 4
0.4	20-01-2016	A. Stasiewicz	Draft	Section 3 and 5
0.5	21-01-2016	L. Porwol	Draft	Consolidation and last corrections
0.6	26-01-2016	E. Osagie	Draft	Updated section 2
0.7	27-01-2016	E. Osagie	Draft	Updated section 6, Section 7
0.8	27-01-2016	A. Stasiewicz	Draft	General updates
0.9	28-01-2016	L. Porwol	Draft	General updates and corrections
1.0	29-01-2016	A. Ojo	Final	Final updates

## TABLE OF CONTENTS

1	Executive Summary .....	7
2	Introduction .....	8
3	Requirements .....	9
3.1	Functional Requirements.....	9
3.2	System Requirements .....	14
4	TET Architecture and Design.....	19
4.1	Architecture Principles.....	19
4.2	TET Reference Architecture .....	19
4.3	Architectural Layers .....	20
4.4	Components design .....	21
5	Implementation .....	26
5.1	Approach.....	26
5.2	Implementation Technology .....	27
5.3	Implementation Details .....	28
5.4	Implemented Features.....	30
5.5	Integration Between TET and SPOD.....	36
6	Deployment .....	37
6.1	Deployment requirements.....	37
6.2	Deployment Process .....	38
6.3	Deployment Issues.....	39
7	Verification and Validation.....	40
7.1	System Verification .....	40
7.2	Testing Framework.....	41
7.3	Test Approach (Agile Methodology) .....	44
7.4	Test Plan .....	45
7.5	System Validation.....	56
7.6	Distinction between Verification and Validation .....	57
7.7	User guide – TET Components.....	57
8	Conclusion .....	70
9	Appendices .....	71
	Test Specification (test cases) .....	71
	Bibliography .....	85

## List of Figures

Figure 1: Scope of ROUTE-TO-PA Requirements.....	9
Figure 2: Mainly TET-Specific Use Case Diagram .....	13
Figure 3. TET Reference Architecture .....	20
Figure 4_TET implemented architecture .....	24
Figure 5 Requirements gathering .....	26
Figure 6 TET main user interface .....	29
Figure 10 Datasets linking.....	33
Figure 11 Detailed User Profile .....	33
Figure 13 Recommendation for related datasets .....	35
Figure 15 SPOD visualisation embedded on TET .....	36
Figure 16 TET OpenID login box.....	36
Figure 17: Validation and Verification Processes in Software Development (Baresi, L & Pezze, M. 2006) .....	40
Figure 18: Considerations for Test Plan, Code Development & Testing .....	41
Figure 19: A systematic approach to system testing. Source: (Baresi & Pezze, 2006) .....	42
Figure 20: The V-Model of Software Development. ....	43
Figure 21: Agile Model - Adopted from ISTQB examcertification.com .....	44
Figure 22: A) V&V Scope, B) V&V Interconnections. (Easterbrook, S. (2010)).....	56
Figure 23: Check metadata quality and check metadata completeness - Screenshot 1 .....	59
Figure 24: Check metadata quality and check metadata completeness - Screenshot 2 .....	59
Figure 25: Profile-based Personalization (user not logged in) - Screenshot 1 .....	60
Figure 26: Profile-based Personalization (user not logged in) - Screenshot 2 .....	61
Figure 27: Profile-based Personalization (user not logged in) - Screenshot 3 .....	61
Figure 28: Personalized search (user login: admin responsibility) - screenshot 1 .....	62
Figure 29: Personalized search (user login: admin responsibility) - screenshot 2 .....	62
Figure 30: Personalized search (user is logged in) - screenshot 1 .....	63
Figure 31: Personalized search (user is logged in) - screenshot 2 .....	64
Figure 32: Personalized search (user is logged in) - screenshot 3 .....	64
Figure 33: Personalized recommendations - screenshot 1 .....	65
Figure 34: Personalized recommendations - screenshot 2 .....	65
Figure 35: : Personalized recommendations - screenshot 3 .....	66
Figure 36: Recommend related datasets - screenshot 1 .....	67
Figure 37: Recommend related datasets - screenshot 2 .....	67
Figure 38: Link dataset - 1.....	68
Figure 39: Link dataset - 2.....	69
Figure 40: Link dataset - 3.....	69

## List of tables

Table 1: Overview of the Societal Activity model of Open Data use .....	11
Table 2: List of ROUTE-TO-PA Use Cases showing the groups of the various cases.....	12
Table 3: Implemented Use Cases Supporting TET Functionalities in Alpha Version or ROUTE-TO-PA.....	13
Table 4: Complete System (Solution) Requirements - TET (implemented components only) .....	14
Table 4: Software Requirements.....	37
Table 5: Hardware Requirements .....	37
Table 6: Links to the repositories.....	38
Table 8: Components for Development and Testing (subject to changes if necessary).....	47
Table 9: Test Data for Alpha Version of TET Components.....	50
Table 10: Test Deliverables .....	51
Table 11: Test Deliverables .....	52
Table 12: Other Testing Tasks.....	54
Table 13: Development and Testing Phase Responsibilities .....	54
Table 14: Rough Schedule of Testing Phases .....	55
Table 15: ROUTE-TO-PA Risk Matrix .....	55
Table 16: TET User Survey - Test 1 .....	60
Table 17: TET User Survey - Test 1 .....	61
Table 18: TET User Survey - Test 1 .....	62
Table 19: TET User Survey - Test 1 .....	63
Table 20: TET User Survey - Test 1 .....	64
Table 21: TET User Survey - Test 1 .....	66
Table 22: TET User Survey - Test 1 .....	67
Table 23: TET User Survey - Test 1 .....	69

## Acronyms

The following are definitions and acronyms used within the document

Term	Definition
OD	Open Data
CC	Creative Common
SPOD	Social Platform of Open Data
TET	Transparency-Enhancing Tools
ICT	Information Communication Technology
DEEP	DataEt-Ecosystem Provider
HTML	HyperText Markup Language
CSS	Cascading Style Sheets,
RRI	Responsible Research and Innovation
GPL	General Public License
PHP	Hypertext Preprocessor
MVC	Model View Controller
LAMP	Linux, Apache, MySQL, PHP/Python/Perl
WAMP	Windows, Apache, MySQL, PHP
IE	Internet Explore
API	Application Program Interface
DOM	Document Object Model
WP	Work Package
RAS	ROUTE-TO-PA Authentication Server
WC	Work Cycle
ODE	Open data Enabler
DoW	Description of Work (i.e. Annex 2 of the Grant Agreement)”
WF	Web Framework
WAF	Web Application Framework
ASD	Agile Software Development

# 1 EXECUTIVE SUMMARY

---

ROUTE-TO-PA (*Raising Open and User-friendly Transparency-Enabling Technologies for Public Administration*) is an innovation project that aims at improving Transparency through prototyping and piloting the integration of Open Data platforms with social networking technologies. The integrated system comprises tools to facilitate improved access and understanding of Open Data to meet information, social, collaboration and decision needs of end users. The two main objectives of the ROUTE-TO-PA project can be summed up as:

1. To enable the transition into the next generation of Open Data portals by creating tools that will enable citizens to engage themselves socially over Open Data resources towards better understanding and more effective use of the available. These tools are collectively termed as *Social Platform for Open Data (SPOD)*.
2. To provide tools that could be integrated into existing Open Data platforms to deliver greater data transparency, quality and understandability. This component is termed the *Transparency Enhancing Toolset (TET)* and it provides basic analytical tools for reducing datasets into a more understandable form for users.

This report is a documentation of activities and results of Task 4.2 of Workpackage WP4 (Technological Development and Integration), carried out between months M7-M12 of the project. The goal of Task 4.2 is to develop the Alpha Version of the Transparency Enhancing Toolset (TET). The report describes in details: 1) the implemented Use Cases and Systems Requirements, 2) TET Architecture, 3) TET Implementation and Deployment models and 5) Verification and Validation process for TET.

TET is implemented as a set of plugins to extend the available features on the well-known CKAN Open Data Platform. The Alpha version of TET described in this report supports the following eight extensions: 1) Support for the use of the WordPress Content Management System as a rich client for CKAN, 2) Enhanced metadata schema to support provenance and alignment with latest W3C guidelines for publishing data on the web, 3) Evaluation of metadata quality for datasets, 4) Linking of related datasets, 5) Enhanced user profiles for personalisation and recommendation, 6) Personalised search and dataset recommendation to users, 7) Recommendation of related dataset to users and 8) Extension of data analytics function on CKAN to support pivot operations on datasets. The development of the above features was carried out through an Agile Software Development process which enabled the development an early prototype and subsequent short “develop-test-release” cycles to engage potential end-users of TET platform.

The target audience for this document includes platform designers, developers as well as various platform stakeholders, including the external testers (non-consortium members) who will participate in the development and testing phase of the project.

## 2 INTRODUCTION

---

As mentioned in the executive summary, this document serves many purposes by assembling many reports from a wide scope of tasks covering:

- the analysed user stories which were presented in the form of use case models and diagrams – **Requirements Specification**. These requirements have been verified against the required standards of quality that will enable them to be used in solution design;
- the system **Architecture and Design** of the selected solution components so far implemented in the *Alpha Version* of the ROUTE-TO-PA project;
- the documentation describing the **System Implementation** approach, technology, model and processes adopted for the **TET** subsystem of ROUTE-TO-PA system;
- the **Deployment** phase literature dealing with the environment and the processes of deployments, elaborating on an instance of deployment with its experiences and issues encountered; and finally
- the **System Verification and Validation** processes which deal with the aspects of testing – beginning from testing framework, test plan, test specifications and then the user guide for system validation by the external users

Each of the itemised task groups is treated under different sections in this document and are further broken down to facilitate in-depth explanations necessary for better understanding of concepts, frameworks, approaches, technologies and processes as they support the innovative technology offered by ROUTE-TO-PA project and as they are employed so far in this project. It suffices to reiterate the fact that this deliverable, although makes mention of the other components, aspects, stages or phases of the ROUTE-TO-PA project. Nevertheless, the focus is on the TET components that are already implemented.

### **Requirement Specifications**

Section two of this document elaborates on the various groups of the requirements of the project – the user or the functional (solution) requirements and the solution quality (non-functional) requirements.

### **Architecture and Design**

Section three documents the TET Architecture, its components and relations between them in order to present the requirements for Transparency-Enhancing Toolset.

### **System Implementation**

Section four describes the implementation of TET: the methodology, technologies and implementation details.

### **Deployment**

Section five defines the system requirements for TET, its deployment process and discusses the issues that may be encountered during the deployment process.

### **System Verification and Validation**

Finally, section six documents the verification and validation process. Components designed and implemented are verified to ensure they met the expected qualities of the system (ROUTE-TO-PA solution) and validated to ensure they meet the user requirements.

### **Conclusion**

Finally, section seven concludes the deliverable findings.



### 3 REQUIREMENTS

---

Guided by the standards of business analysis for IT product development, ROUTE-TO-PA solution requirements were gathered from interviews with a wide range of stakeholders and requirement analysis workshops in multiple locations across the EU. The requirements were enriched with detailed investigations and analyses of relevant knowledge sources including online and academic publications. The scope of the requirements is visualised in Figure 1.

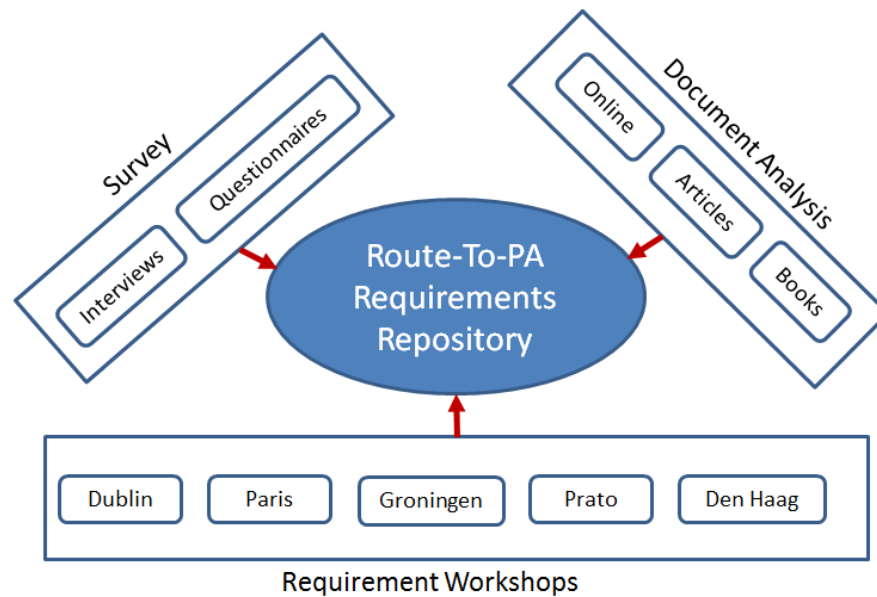


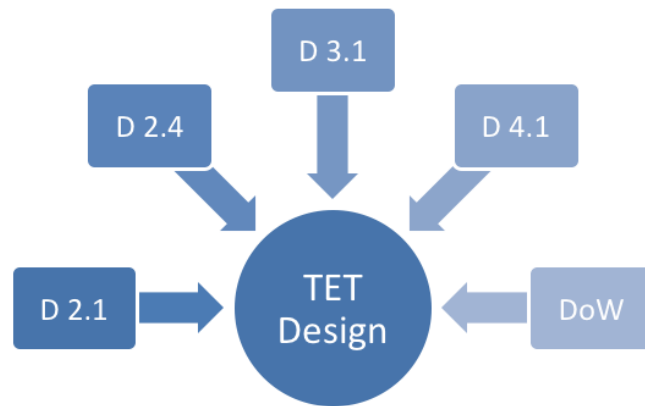
Figure 1: Scope of ROUTE-TO-PA Requirements

#### 3.1 FUNCTIONAL REQUIREMENTS

##### 3.1.1 OVERVIEW OF REQUIREMENTS

The user stories and gathered requirements became a base for a comprehensive set of use cases, essential for the reference model specification.

The use cases and models developed based on the requirements from field work were presented in deliverable *D2.4: Requirements Specification and Use Case Models*. Earlier deliverables include *Deliverables D2.2: Analytical Framework and Initial Scenarios on Open Data and Transparency*; and *Deliverable D2.3: User Stories on Open Data and Transparency*. These deliverables gave detail accounts of the user scenarios and user stories from which the Requirement Specifications and Use Case Models were developed [IBM.com; Coursehero.com].



*Figure 2 Requirements gathering*

Requirements for TET were collected by analyzing the project deliverables: D2.1, D2.4, D3.1 and D4.1 as well as DoW. For the purpose of better requirements traceability, the earlier tasks of the project including requirements elicitation and analysis, and their respective documentation were planned and executed to align with ROUTE-TO-PA objectives and the democratic values they demonstrate or exemplify. These objectives include:

- Integration of related datasets using Linked Data principles
- To provide basic analytics on datasets to detect violations of rules (e.g. transparency related rules), that is, where metadata supply for the purpose of enhancing data searchability and linking is lacking through the metadata completeness measurement.
- Personalisation – exploitation of social media profiles of users, e.g. demographic features and interests, the possibility of analysing contents contributed on social media (e.g. topic detection).
- Dataset Quality – generate reports on quality of metadata and provenance information (e.g. based on specific standards) – through view metadata functionality
- Recommendation – clustering of datasets based on categories, and usage patterns (e.g. datasets downloaded or requested together) as popular searches and consumption pattern or as natural categories e.g. budget datasets, population datasets, economic datasets, etc.

These objectives were mapped to see how they demonstrate democratic values: Monitorial Democracy, Deliberative Democracy and Participatory Democracy in the earlier documentation referred to above and modified in Table 1.

Table 1: Overview of the Societal Activity model of Open Data use

Use of open data	Monitorial democracy Monitoring government behavior	Deliberative democracy Feeding public debates	Participatory democracy Enabling collective action
<b>Object</b>	Policy issue	Policy issue	Policy issue
<b>Subject Citizen</b>	Watchdog	Partner in dialogue	Partner in Action
<b>Tool</b>	<ul style="list-style-type: none"> <li>- TET ensures transparency by access to data, data search and tools to visualize the data</li> <li>- SPOD facilitates data sharing</li> </ul>	<ul style="list-style-type: none"> <li>- TET ensures access and enables data analysis, personalization, comparing and combining datasets and feedback options</li> <li>- SPOD facilitates participation with debate and feedback options</li> </ul>	<ul style="list-style-type: none"> <li>- TET ensures access, decision-making tools, the ability to add data</li> <li>- SPOD facilitates collaboration with government and other stakeholders e.g. refining solutions and working together in documents</li> </ul>
<b>Rules/Mechanism</b>	Transparency, checks and balances, e.g. foia laws, privacy laws, but also rules regarding the virtual community	Deliberation, open communication e.g. procedures regarding participation	Collaboration or co-creation e.g. procedures regarding who is invited (based on expertise)?
<b>Community</b>	Government, journalist	Government and citizens	Government, citizens, businesses, researchers
<b>Role</b> e.g. government	Provider of information	Facilitating dialogue	Partner in Action
<b>Outcome</b>	Critical view on government behavior	Contribution to debate about policy issue	(Collective) action to produce public value

### 3.1.2 USE CASE MODELS

**Functional Requirements:** We presented the scope of *Functional (User) Requirement Specifications* and alignment with relevant project objectives and democratic values. However, in this deliverable, we intend to focus on the use case specifications and models that are TET-related. Furthermore, the current version of TET – Alpha version deals only with the components of TET that have been developed (i.e. implemented). As the remaining use cases are implemented in iterations of Agile approach, they will be released in subsequent versions of the TET subsystem along with the necessary documentation.

To deliver an overview of the use cases and the models, the team broke down the entire use case collections into three groups. The first group details the TET-related components (TET functionalities for TET objectives),

the second group details the SPOD-related components (SPOD functionalities for SPOD objectives) and the third group details the functionalities common to both TET and SPOD objectives. The partitioning of the use case models is well presented in deliverable *D2.4: Requirements Specification and Use Case Models*. Considering the fact that this documentation is focused on the Alpha version of the implemented TET components, which support TET specific objectives of the ROUTE-TO-PA project, there is no emphasis placed on the other two groups of the requirements specification seen in Table 2.

Table 2: List of ROUTE-TO-PA Use Cases showing the groups of the various cases

Use Case ID	Group	Primary Actor	Use Cases	U/Story ID
UC1	TET	User (Entrepreneur)	Check Metadata Completeness	S29.1
UC2	TET	User (Entrepreneur)	Add/Update/View Provenance and metadata-related to a Dataset	S29.3
UC4(A)	TET / SPOD	User (Resident)	Provide Supporting Content on SPOD Post	S2.3
UC4(B)	TET / SPOD	User (Resident)	Share a dataset with a visualisation on SPOD	S2.1
UC5	TET / SPOD	User (PA)	Provide Supporting Content on SPOD Post (Share Link to Dataset/File via SPOD Post)	S35.1
UC6	TET	User (Community Activist)	Take a Screenshot and Annotate it	S7.1
UC7	SPOD	User (Entrepreneur)	Attach Widgets and Tags Tools on SPOD Forum	S31.1
UC8	TET / SPOD	User (Entrepreneur)	Using External Social Media for Login Authentication: e.g. Facebook (External Social Media Authentication)	S3.0
UC9	SPOD	User (Network Member)	Voting on SPOD FORUM	S2.6
UC10	TET	User (Resident)	Analyse a Dataset	S1.4
UC11	TET	User (PA)	Query Data	S32.2
UC12	TET	User (Community Activist)	Integrate a Dataset	S8.1
UC14	TET	User (Community Activist)	Link a Dataset	S8.1
UC15	TET	User (Student leader)	Add a Dataset to Compare	S13.9
UC16	TET	User (Student leader)	Compare Datasets	S13.9
UC17	TET	User (Consultant)	Personalize Search	S28.2
UC18	SPOD	User (Entrepreneur)	Receive Notification for New Datasets	S3.1
UC19	TET/SPOD	User (Consultant)	Request Recommendations for Data Sets	S26.5
UC20	TET	User (Resident)	View a Chart	S2.2
UC21	TET	User (Resident)	View Infographic	S2.2
UC22	SPOD	User (Consultant)	Post on SPOD	S28.3
UC23	TET/SPOD	User (Consultant)	Login	S28.1
UC24	TET	User (Entrepreneur)	Enrich Profile	S38.1
UC26	SPOD	User (Citizen/PA)	Add an Event	
UC27	SPOD	User (Community	Create a New Topic Enabling an Anonymous	



UC23	TET/SPOD	User (Consultant)	User Login	
UC19	TET/SPOD	User (Consultant)	Request Recommendations for Data Sets	S26.5
UC24	TET	User (Entrepreneur)	Enrich Profile	S38.1

## 3.2 SYSTEM REQUIREMENTS

### 3.2.1 OVERVIEW OF SYSTEM REQUIREMENTS

The use cases listed are derived from a consideration of user stories and the breakdown of the specific and unique functionalities and capabilities that the system must offer to meet the needs of users expressed in the user stories that were provided by stakeholders. However, these cases are only a fraction of the requirements – the **Functional Requirements** portion of the solution (system) which needs to be developed [(Armin B. (n.d)); Pamela V. L. (2013)]. To create a complete set of system (or solution) requirements, the functional requirements need to be supplemented with and supported by another set of **Quality Specifications (quality requirements)**. These groups of requirements in terms of software engineering define the performance, the speed, the usability, the navigability and the reliability qualities, etc., of the software. Details of this topic are discussed below in section 3.2.2 below.

For the purpose of the ROUTE-TO-PA project, the quality requirements are selected in such a way that they go beyond the IT product being developed. The qualities must enable the system to create both user-requested values (e.g. accessibility to datasets, view metadata quality of datasets, link datasets with the efficiency of performance) and the democratic values embedded in TET concept at the same time. This means that TET subsystem quality parameters and metrics (qualitative or quantitative) must justify or support data transparency and accessibility to datasets through capabilities that enable data consumers and citizens to understand datasets easily, gain access to the datasets, analyse the datasets and seamlessly receive datasets that meet their interests in terms of relevance to them, quality and usability. In the table below, we supply the functional requirements for the components already implemented and match each of those requirements with the system qualities that have been used to upgrade the functional requirements up to the system requirements level so as to enable the implemented solution to meet the needs of the stakeholders.

Table 4: Complete System (Solution) Requirements - TET (implemented components only)

Use Case	Functional Requirements	TET Objectives	System Quality Requirements
UC23	User Login	The system provides login form data fields for you to complete and save to user account and to enable the user to enjoy a better experience with the platform.	Availability, Operability, Portability, Data Security and Privacy, Simplicity
UC1	Check Metadata Completeness	Visualise metadata completeness quality for a dataset by displaying the measurement bar, showing percentage rating for the completed metadata fields for the dataset to enable users to decide on the suitability for use or fit-for-purpose	Usability, Simplicity, Integrity, Accessibility, Accuracy
UC2	Add/Update/View Provenance and metadata-related to a dataset	System support for provenance records: values entered for the – source, versions (current & previous) and frequency of publication fields in the metadata screen of the dataset.	Usability, Operability, Auditability, Integrity and Auditability
UC10	Analyse a Dataset	System provides simple analysis, analytics and visualisation tools to enable users (including non-tech savvy users) to	Usability, Supportability, Simplicity, Clarity,

		analyse datasets & visualise results in common graphs to enhance comprehension of the data meaning	Correctness, Intuitiveness, Performance, Accuracy
UC14	Link a Dataset	The part of this objective so far implemented is the part that enables dataset linking for searchability and recommendation purposes as support for dataset accessibility. Dataset linking for the purpose of dataset integration has not been implemented.	Supportability (file formats), Usability, Operability, Extensibility, Traceability and Accessibility.
UC17	Personalize Search (User not logged in)	The system provides options e.g. role, location, dataset category, etc., for the user to personalise search by default. This supports dataset accessibility and improved search experience for users	Operability, Usability, Reliability, Accuracy, Data security & Privacy, Performance
UC17	Personalize Search (User logged in)	The system provides options e.g. interest, role, location, dataset category, etc., for the user to personalise search as system support for dataset accessibility and improved search experience for users.	Operability, Usability, Reliability, Accuracy, Data security & Privacy, Simplicity, Performance
UC19	Recommendations for Datasets	Provide the user with dataset recommendation options in the form of Dataset category or subject or by popular searches to enable user search within recommended options while not logged in.	Usability, Personalisation, Simplicity / Interface friendliness, Accuracy and Performance
UC19	Request Recommendations for Datasets	Provide the user with options (location, role, age, email, sex, marital status, disability, interest, etc.) to enable user personalise account & receive content suggestions based on account data. <i>(Not implemented yet: option for user to opt in for data recommendation alert via email)</i>	Usability, Personalisation, Simplicity / Interface friendliness, Data Privacy & Security, Integrity and Performance
UC24	Enrich Profile	System provides more personal data options e.g. location, age, marital status, sex to enable users adds more data to profile so that user can search & receive more relevant resources based on his/her enriched profile details	Operability, Accuracy, Accessibility, Dependability, Data Privacy & Security

### 3.2.2 EXPLANATIONS OF SYSTEM QUALITIES

The product quality requirements also referred to as the **Non-functional Requirements** are a set of standard specifications, which a software product must exhibit when in use by stakeholders. When these specifications are combined with the functional requirements, the latter are transformed into complete **System Requirements – the inputs into the** solutions to the needs of stakeholders to their problems. The system, software or solution requirements are therefore used to design the software product (or the components thereof) that should meet the needs of the various stakeholders according to their desires as captured in the user stories and exemplified in the scenarios (see table Table 4).

#### Accessibility

This quality ensures that the system functionality can be reached by a user at any given time. Even though a system's functionality or tool may be available in its physical computer resource form but if the user cannot, for any reason login into the system or click on a button to invoke the function, it is not yet accessible to the user. Accessibility is closely related to the understandability and availability qualities in that if the system poses a difficulty in assimilating the instructions on how to use a tool or a given functionality is generally misleading, then a user cannot have the advantage of the resource. In this case, this functionality is neither accessible nor available to the user. Furthermore, the accessibility quality also applies to the quality of being able to support

adaptability for use by people of special considerations such as people with disabilities. Furthermore, accessibility should enable the user to use the resource to reach the dataset of relevance to him or her thus supporting the dataset accessibility as a tool for promoting transparency in the context of ROUTE-TO-PA project.

#### **Accuracy**

The system should work as expected and produce the result that is expected. Thus, a user should find the system in operation according to how it is intended to function and produce the expected output. Clicking on the login button after entering the correct login details, the system should accurately login the user to the landing page assuming all other constraining factors such as poor Internet connection are taken care off. Accuracy also means that the system will produce output according to the input parameters meant for the system to use in fetching certain resources or outputs. For Example, the results of all mathematical computations should be error-free, and the system should utilize search parameters to search and retrieve the required datasets if they are available on the portal with 100% accuracy.

#### **Auditability**

This quality relates more to the capability of TET tool to permit an efficiency check rather on the data transparency feature it tends to support than the working capability of the functions itself. Its auditability feature will permit a systematic check or assessment of the open data practice or standards it supports to allow a user understand the accuracy, efficiency or effectiveness of the dataset or other open data resource operations it supports. Such understanding may help user decide on certain issues pertaining to use of a given resource on the platform

#### **Availability**

Availability quality, as explained under accessibility means that the feature or tool or functionality of the system must be there for the user to use at all times. The system must not leave the user in any form of frustration while attempting to use the resources due to the absence of a tool. The quality of system availability extends to the capability to keep uptime in good measure – that is high – over a given period of time. Conversely, the downtime should be low – meaning the system functions remain in ready-for-use all the time it is needed by a user, all things being equal.

#### **Data Privacy and Security**

In cases of login with personal details, the systems must adhere to the required standard of maintaining the safety of user personal details and maintain the user account and profile data against unauthorised access.

#### **Integrity**

The system should maintain a level of standard that can be adhered to as moral principles of completeness and wholeness in output and the quality thereof as stated governing professional or occupational standards in the domain.

#### **Interface / User friendliness**

The system interface is the point of interaction between the system and the user or another system. On the human user interface point, the system should present an attractive easy to engage windows containing screen features that are not eye-straining. Generally, the interface should adhere to ergonomic standards and contain readable fonts that the user would always like to return to at any time. The aesthetics should be that which is superb in quality with attractive colour mix differentiating the buttons into groups of functions. Similarly, the interface between systems should permit interaction and ‘hand-shaking’ between the element of both systems in system engagements that enable interoperability, data sharing or other synchronisation operations.

#### **Operability**



The buttons and tools implemented must be deplorable in the manner intended to allow the user use them to achieve the desired goals. Thus, there should be no unexpected errors or hitches capable of forestalling the appropriate interaction of the user with the system. System tools should be able to support the given operations for which the tools and functions were implemented to execute seamlessly in order that users can utilise the system as an application software to achieve a goal meant to be done with the system.

#### **Performance**

Performance quality relates to the capability of the system to execute operations with high efficiency and speed. Thus, the system speed should be up to the reasonably expected short time to execute a function e.g. return a query result after search criteria have been entered and the search button pressed. Performance quality is also defined by the smoothness with which interactions as a result of input and return operations are executed throughout the process involving the activities to use the application to achieve a goal by a user.

#### **Reliability**

Reliability quality calls for the capability for the system to maintain the expectations that a user places on the system. Reliability, therefore, relates to the factors like performance, availability or meeting up to the standard of delivering the query results with accuracy. It means that the system should be able to do what it should all the time when a user invokes the functions under different environments that it is meant to function well.

#### **Simplicity**

Simplicity is the quality that ensures that users are able to learn the application and use it without much stress in understanding how to apply the tools. User friendliness of the interface may be enhanced by the quality of simplicity which avoids unnecessary complexity in the interface or windows of the application. ***A simple system is one which is not complex and hence may permit users to learn how to use the system without a profound training process. Simplicity can be enhanced with tooltips and other system documentation available to users on the system windows without causing cluttering.***

#### **Usability**

Usability quality means the implemented components have the quality of being able to be applied for a specific purpose by the user. It is important that this quality is maintained for all environments and platforms in which the system has been stated to be supported. If the system is not usable, it automatically loses many of the other qualities because a system has to be usable before the need to talk about other qualities of the system. If a system is not usable, it means that it cannot be deployed to achieve a given purpose; and this can be due to many reasons including very serious ones such a faulty requirement specification right from the elicitation process or code development or system architectural blunder.

#### **Supportability and Inter-operability**

A system will surely operate with other systems, and it is hard to find a standalone system that does not communicate with other systems or platforms. To operate smoothly with other systems, components and interfaces, a system must maintain the capability to utilise standardised file formats, synchronised data components and other required system protocols. Possession of these qualities will enable the system to have a good 'hand-shaking' with other system or inter-operate with other applications smoothly. This capability will enable the users to get the best of advantage from the given system.

### **3.2.3 OTHER REQUIREMENTS (TRANSITION REQUIREMENTS)**

The other group of requirements otherwise known as the **Transition Requirements** are the group of requirements that are needed to support the periods of replacement of an old version of software with a newer version of it. These requirements are needed only while the transition period is in progress; hence the name *Transition*. Once the complete replacement of the old with the new system is completed, the

transition requirements become redundant and are then discarded. The function of the requirements is mainly to support data migration from the old to the new systems, support the interfacing of the two systems while they are being run simultaneously if the business or operation of the organisation cannot be halted during the fading out period of the old version. In the case of the new software product offered by ROUTE-TO-PA project, there is, at this stage, no data migration operations involved in this software system implementation as it is a new version. However, other stakeholders who might be deploying the systems as an application system in their operations may have the need to define more detailed Transition Requirements in due course. Again, components of Transition requirements such recruitment of skills, training of personnel, education for users and production of user guides are relevant to the various versions of this system as they are released. Each release will be supported by the aforementioned components by the Insight Centre with respect to TET subsystem.

#### **User Guides**

This is a textual material that provides instructions on how a user can operate the system to achieve a specified goal. The guide is written for each component of the implemented TET functionality to test and validate the system and to familiarise users with the new platform. The aim is also to enable the users to report their findings as well as their satisfaction with the implemented functionalities in line with the use case models and the user stories. In order to facilitate user rating, remarks and critical scrutiny of the new system and to collect relevant metrics on implementation qualities as well as the levels of satisfaction, a few survey questions are attached to each component of TET in the user guide. More on the user guide is provided in Executive Summary of this document.

#### **Skills Education and Training Needs**

These are a closely related set of transition requirements and are more relevant to the user end of the deployment process than the development team end of it. However, on the Development/BA team side, we have put in place continuous training process and application of relevant (properly referenced) concepts, ideas and templates from relevant organisations in order to gain skills to implement the components and pre-test them internally. On the user side, each user as an individual or organisation will have to provide the required resources to bridge the skill gap by upgrading to the required skill level. To facilitate the closure of the skills gap, the Insight Centre Development/BA team has produced a comprehensive User Guide for this purpose. A better more detailed User Guide will be produced for ROUTE-TO-PA product as a deliverable on its own towards the end of the project. Lastly, the Development/BA team has also specified the environment required for testing the implemented components for the Testers to have in place for the testing exercises.

## 4 TET ARCHITECTURE AND DESIGN

---

This section documents the TET Architecture, its components and relations between created accordingly to the requirements for Transparency-Enhancing Toolset (TET).

### 4.1 ARCHITECTURE PRINCIPLES

In order to understand the requirements for Transparency-Enhancing Toolset (TET) Architecture we analyzed a selection of existing Open Data Platforms and followed Use Stories described in D2.3 (ref. *D2.3 User stories on Open Data and Transparency*). The selection of Open Data Platforms was made by investigation of the latest publications concerning Linked Open Data and of the usage of the available platforms.

Moreover based on publicly accessible documentation (projects websites, documentation, publications, press releases) each platform was analyzed in order to investigate the key components and relations between them. The patterns observed across the reviewed Open Data Platforms, and research results were used to create the TET Architecture.

Architecture is defined as *the fundamental organization of a system embodied in its components, their relationships to each other, and to the environment, and the principles guiding its design and evolution* [IEEE 1471, 2000]. The purpose of architecture is to guide the development and the implementations of the platform. Moreover, it allows understanding of the requirements, the workflow and the challenges faced during the deployment of the solution.

### 4.2 TET REFERENCE ARCHITECTURE

Generic TET reference architecture is presented in Figure 4, and it is organised in three Architectural Layers: (i) User Interface, (ii) Services, and (iii) Storage.

A detailed description of the layers is described in section 4.3 while the components details are explained in section 4.4.

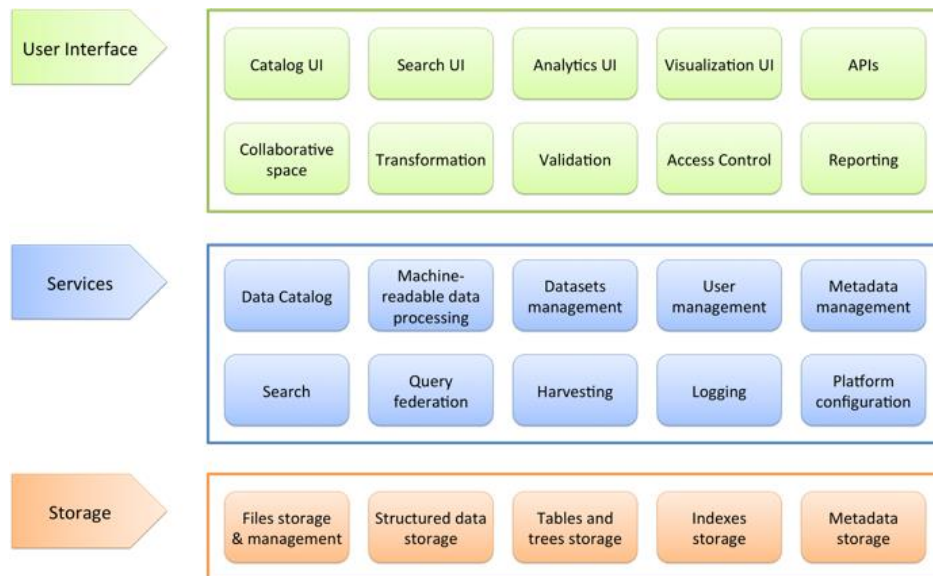


Figure 4. TET Reference Architecture

### 4.3 ARCHITECTURAL LAYERS

Different layers in the software architecture are allocated different responsibilities. In our case, the layer at the bottom handles the data. The middle layer can be viewed as a layer of the high level of services provided on top the data while the top layer provides services to users and applications. Details of each layer are as follows:

#### User Interface layer

The User Interface layer provides basic portal functions such as access to the data, search interface, personalization and customization features, etc. The search feature allows users to find quickly information stored in the portal, while analysis and visualizations features allow users to explore, analyse and visualize various types of data, such as tabular and geospatial data. Various APIs allows external applications to consume services offered by the platform.

#### Services layer

The Services layer provides services on top of Storage layer that can be exploited by the User Interface layer. Data Catalog services are used to list the details of datasets and associated metadata stored in the platform. Search service uses the search index to relevant content. Platform extensions services allow external applications to use the platform services. All these services have the corresponding features in the interface layer.


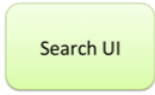
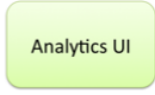
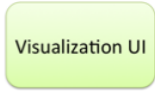
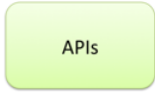

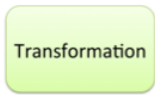
#### Storage layer

The Storage layer is concerned with persistence of data and provides services for efficient data storage and retrieval. This layer is responsible for the management of file storage, structured data and search indexes. Data can be stored directly in file system storage or in the structured data store.

#### 4.4 COMPONENTS DESIGN

The detailed structure of the Architecture is shown in the tables below. For each of the layers, the components role and description are as follows:




##### User Interface

Component name	Role / description
 Catalog UI	The user interface allows browsing, exploration and discovery of datasets. Basic functionalities are search, browse, and management of datasets and associated metadata, etc. It provides additional information that facilitates in dataset discovery, e.g., featured datasets, additional labels, suggested datasets, etc.
 Search UI	Search user interface allows full-text search queries over a collection of datasets. Moreover, it provides personalised results based on user profile. Basic functionality is the text search.
 Analytics UI	Comprises a number of approaches to generating new insights from data: data mining, data fusion, spatial linking, statistical analysis, clustering, and so on. Usually, knowledge of these techniques is necessary for a user to derive value from the interface. The alpha version of TET provides Pivot View tool. Next releases will enhance and extended this component by a set of the tool: Aggregation, Machine Learning, OLAP, Text Analytics, Network Analytics and LinkedData Analysis.
 Visualization UI	Allows users to visualise the datasets (i.e. map view, chart, tabular view) as well as the results of the analysis for non-expert users. This interface helps with making data more accessible.
 APIs	Allows external applications to consume the services offered by the platform in easy programmatic access. In the first release of TET, the main focus was to provide various statistics about TET usage.
 Collaborative space	Displays user profiles; user registration; permission management. Basic functionalities allow users to communicate (messaging system) and discuss the datasets.
 Transformation	Allows export of data in additional formats. An example is XML to CSV conversion. The future releases of TET it will support the data refinement and cleansing.

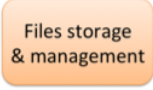

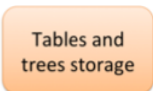


Validation	Syntactical validation of raw data files for selected raw data formats, e.g., XML parsing, CSV parsing. It should gracefully handle i.e. syntax errors.
Access Control	Access Control Logic and restricts access to some parts of the system that needs to be regulated.
Reporting	Provides reporting statistics, data summary, queries summary, datasets quality summary and so on.



Component name	Role / description
Data Catalog	Data Catalog allows browsing, exploration and querying of a collection of dataset metadata records.
Machine-readable data processing	Processes and serves available data in a format that can be easily understood and consumed by a machine.
Datasets management	Handles dataset management related issue i.e. creation, editing, storage, etc.
User management	Stores and manages data related to the user. Provides user the ability to register, edit their profile, and view profiles of other users. Moreover, this component controls the access, permissions and group membership.
Metadata management	This component allows creation and editing of metadata records, such as provenance information, modification date, license and so on. Metadata management also includes quality assessment.
Search	This component is responsible for the retrieval of information according to the user queries and actions. It is directly connected with the data stores. It allows full-text search over stored data, at sufficiently fine granularity. This includes metadata search, and searches all the data, potentially returning any resource. TET provides personalised search based on user profile: role and category.
Query federation	Query federation allows data integration between multiple instances of the data portal.

	Harvesting services allow integration of data from other portal instances and other, external data sources.
	This component provides logging services throughout the system. All components should log to a single system in a consistent way. This component helps to monitor the system status and issue appropriate alerts in the case of system failures.
	This component allows modifying the platform parameters (such as timeouts, limitations) and the available functionalities. In alpha version of TET, it was implemented through a configuration file (enabling extensions, system paths) and backend panel (content).



Component name	Role / description
	The main role of this component is persistent storage of raw data files and the dynamic content of the sub-pages. It is implemented as a local datastore.
	Structured datastore allows efficient retrieval and querying data. File and records stored on the platform can be indexed for efficient retrieval for searching.
	Indexed and structured storage of tables and trees. TET solution is based on SQL database.
	This component is responsible for storage of various schema elements, search indexes, vocabularies, concept schemes, etc.
	This component includes various kinds of persistent metadata, including provenance information.

#### 4.4.1 IMPLEMENTATION ARCHITECTURE

TET implementation architecture comprises of Frontend, Data Platform and TET Analytics. The frontend is used to handle interaction with the user and provides neat content management capabilities. Data platform handles data management related issues while TET Analytics is used to offer advance data analysis capabilities over a large volume of data. The architecture is presented in **Errore. L'origine riferimento non è stata trovata..**

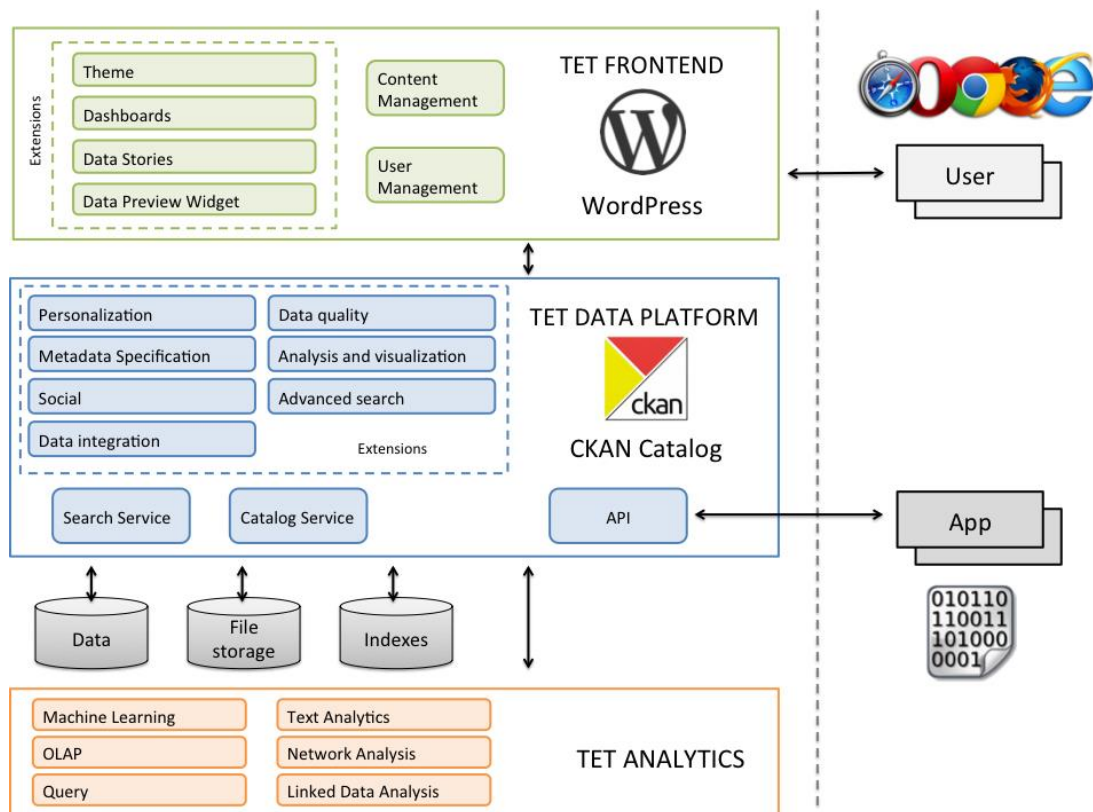


Figure 5 TET implemented architecture

### Frontend

WordPress, content management system, is used as front-end for TET toolkit suite and is extended with additional plugins to support dashboards, storytelling and to render different views on data stored in data platform.

### Data Platform

CKAN is used as data management middleware for TET; it provides features such data catalogue, search and management of datasets. Additional extensions are added to CKAN to enable support for personalization, rich metadata specification, analysis and visualization, advance search over structured data, data integration, sharing, and tools for assessing data quality. The layer uses a database, file storage and search index to offer a wide range of services over datasets to users and to Apps via API. Moreover, its role is to store, manage and serve all kinds of data to the user in a unified view. It also serves as user Interface for dataset catalogue.

### TET Analytics

TET analytics platform offers advanced analytics services like Machine Learning, OLAP, Text Analytics, Network Analysis and Linked Data Analysis, etc. to the data platform. Existing analytics tools like R, Python and Spark, etc. will be used to build an intelligent Open Data platform that will enable users to make better use of the Open Data as a key instrument for better transparency and better decision-making.

Extensions provide a link between the User Interface and the Services. It exposes all kind of analytics and visualization services available in TET. Moreover APIs & widgets are providing a bridge between SPOD and TET to keep a unified state of the data between platforms: i.e. user status. Moreover, it allows including functionalities of the two platforms in each other in a seamless way (i.e. link to start the discussion).



End users interact with the application via the user interface provided by WordPress and CKAN. Based on action selected at the user interface either CKAN default behaviour or TET extension are triggered, and depending on action data might be fetched from storage or database and processed by the relevant module, at an appropriate response is rendered on the user interface. External applications can access features provided by the front-end or back-end.

## 5 IMPLEMENTATION

---

This section describes the implementation of TET. In particular, we define the details of the components included in TET, functional and implementation descriptions; the methodology followed as well as the technologies used during the development.

### 5.1 APPROACH

To effectively manage the software development process and to deliver high-quality software, we leveraged the agile development methodology. Agile methodology encourages continuous iteration of development and testing throughout the software development lifecycle [7]. This way the software evolves through communication between self-organizing teams with cross-functional expertise.

Agile methodologies have been proposed by the design science and software engineering communities as a set of flexible techniques that can adapt to the changing of user requirements [1]. The flexibility in software development projects is always important in order to increase the response to requirements changes (e.g., Byrd and Turner 2000, Duncan 1995, Gefen and Keil 1998, Lee and Xia 2005, MacCormack et al. 2001) - the positive relationship between flexible development architecture and software project performance as well as a positive relationship between team flexibility and end-user satisfaction with a system (Lee and Xia (2005). However, flexibility is not without cost and also tightly inimical to the degree of structure embedded in a development process [1].

It has been always the business customer responsibility for the success or failure of the product as a product owner. Hence, some agile technologies including SCRUM define this role by key functions, which include maintaining and prioritizing the product backlog, specifying and gathering individual user stories and their acceptance criteria [2].

Traditional SCRUM centers the product owner's interactions on two main artifacts which are a project plan (which relates the owner to the stakeholders) and a backlog (which relates the owner to the development team), where the product owner plays two roles in the development team but also deal with logistical challenges like customer bug reports for example [2].

SCRUM is defined as a framework within which people can address complex adaptive problems while productively and creatively delivering products of the highest possible value. Where, main features of any SCRUM is lightweight, simple to understand and difficult to master [3].

SCRUM started since the early 1990s to manage complex product developments. It deals not only with the process but also the framework where you can employ many techniques and processes. It binds events, roles and artifacts as well as managing the relationships and interaction between them [3]. SCRUM team contains the product owner, development team and SCRUM master. However Scrum events include the sprint, sprint planning, daily scrum, sprint review and finally sprint retrospective [3].

We adopted SCRUM to manage product development. Tasks are divided into sprints and useable software with incremental changes is delivered at the end each sprint. Planning meetings at the start of each sprint are organized to define clearly the objectives of each sprint and to plan development tasks. To track the project progress; daily short stand-up meetings are held. A retrospective session is conducted at the end of each sprint to learn from the sprint experience and use it for improving future work. Initially, we started with long sprints but gradually will use short sprints. The software is extensively tested before each release. Feedback from the pilots is used to prioritize development tasks, bugfixes and enhance implemented features.

## 5.2 IMPLEMENTATION TECHNOLOGY

TET is set of tools to build on top of two popular platforms WordPress and CKAN. Wordpress is used as front-end and provides extensive content management system capabilities. CKAN is a de facto standard for open data portals; it provides comprehensive data management services.

WordPress is free open source content management system. It's written in PHP and uses MySQL as a database. WordPress supports theming and extensions with plugin mechanism. WordPress has a vibrant community of supporters and developers; there is a huge library of themes and extensions available that are ready to be used with WordPress. Other features include support for mobile devices, search engine optimization, tagging and permalinks. WordPress started as a blogging platform; later on, it changed to the full-fledged content management system (CMS). WordPress can manage every aspect of a website like creation, organization and search engine optimization [4]. Core features support by WordPress can be summarized as following:

- 1) Custom Taxonomies,
- 2) No rebuilding,
- 3) Custom Post Types,
- 4) WordPress Themes,
- 5) Cross-blog communication tools,
- 6) Spam protection,
- 7) Full user registration,
- 8) Password Protected Posts,
- 9) Easy Importing,
- 10) XML-RPC interface,
- 11) Workflow,
- 12) Intelligent text formatting and Menu management [4].

Comprehensive Knowledge Archive Network (CKAN) is the free open source web-based data management solution. It's written in Python and uses Pylons framework. It uses Postgres<sup>1</sup> as database and apache solar for the index to fulltext search. CKAN facilitates easy discovery of data and has a broad user base. Governments around the world use it for serving open data, which includes US, UK, Ireland and Australia. CKAN provides provision for theming and extensions. CKAN provides many essential features required to publish, share and visualize datasets and offers powerful cataloging, searching and storing capabilities [5]. CKAN software development began in March 2006 and first release in July 2007. CKAN project is implemented and managed by Open Knowledge Foundation (OKF) [6]. The main objectives of CKAN can be summarized by:

---

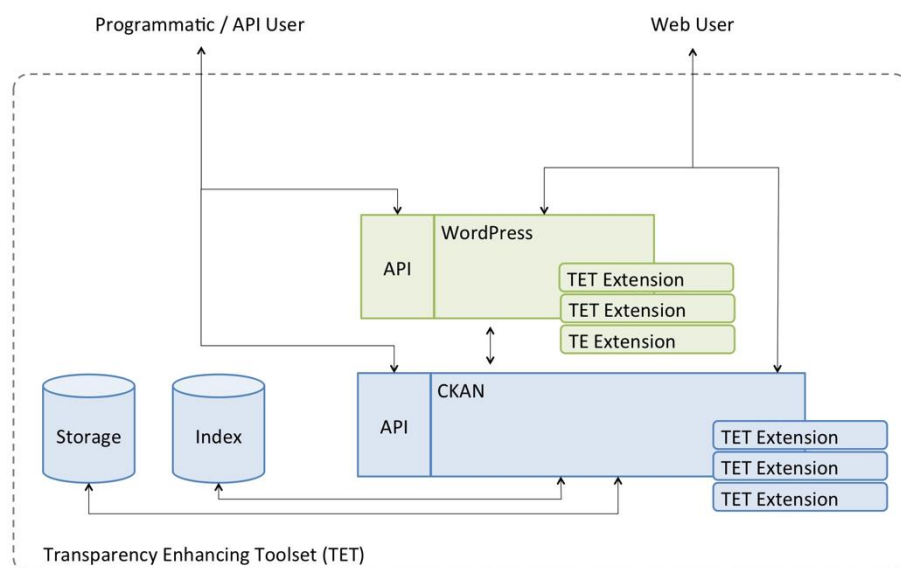
<sup>1</sup> <http://www.postgresql.org/>

- 1) Freedom of access, creation and dissemination of knowledge
- 2) Develop, support and promote projects, communities and tools that foster and facilitate the creation, access to and dissemination of knowledge.
- 3) To campaign against restrictions both legal and non-legal on the creation, access to and dissemination of knowledge.
- 4) Act as an intermediary between funding institutions, the creation and diffusion of knowledge projects [6]

TET exploits theming, and extension capabilities offered by WordPress and CKAN to deliver rich user experience, powerful data discovery and analysis features to help users in leveraging the real potential of open data. TET reuses existing free opensource tools and libraries wherever possible. The majority of TET code is written in Python, PHP and JavaScript. Most of TET functionality relay on API provided by WordPress and CKAN, however, external libraries are used where there is no native support available in WordPress and CKAN. In some cases where there is existing solution already available, we have tried to reuse it.

### 5.3 IMPLEMENTATION DETAILS

TET development is based on two different systems WordPress and CKAN. TET tools are designed and developed as extensions to both the platforms. Figure 5 illustrates how both the platforms are stitched together to come up with unified experiences.



*Figure 5 CKAN and WordPress integration*

Both platforms use different programming languages, tools and libraries; this inevitably makes the task of setting-up of the development very complicated. To reduce the time required to setup the development environment and to have consistent development environments across all development machines, we created a Vagrant script. Vagrant<sup>2</sup> is a tool for creating and distributed development environment. Development environments created with Vagrant can be easily deployed on development machines and on the cloud. TET development activities can be broadly divided into Theming, WordPress extensions and CKAN extensions development.

<sup>2</sup> <https://www.vagrantup.com/>

## Theming

TET uses WordPress platform to add content management system capability on top of CKAN. Both platforms have quite a distinct look and feel. To have a consistent look and feel across both platforms, we created themes for WordPress and CKAN. New color scheme and graphics were added to improve the aesthetics (Figure 6).

New navigation and user focused interface was designed that enables easy discovery of datasets. New widgets were added to the interface that allows useful information from CKAN to be displayed in WordPress. These extensions make use of WordPress API, jQuery<sup>3</sup> and Bootstrap<sup>4</sup>. Theming CKAN is completely different than WordPress; CKAN uses jinja templates and has a different way for theming, which in some ways is quite restrictive and challenging task.

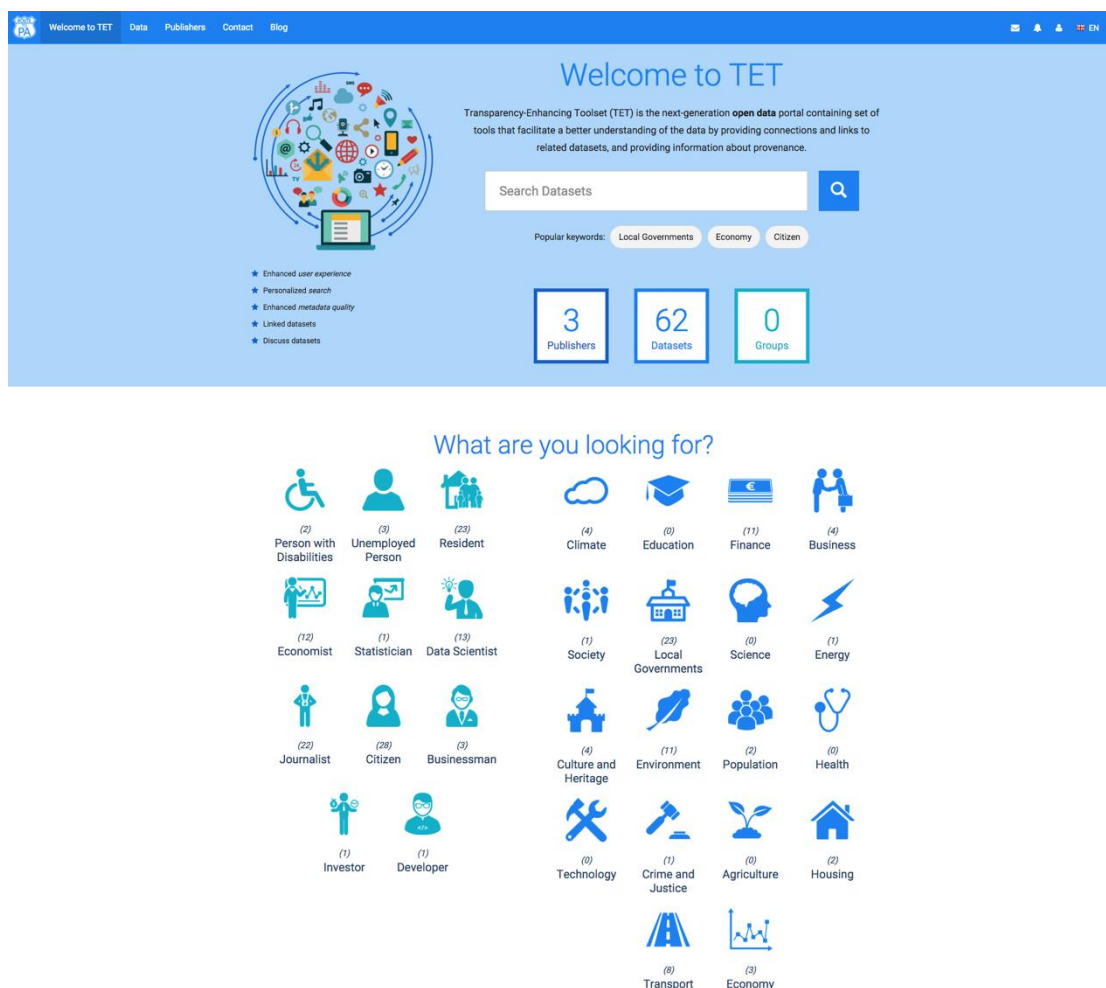


Figure 6 TET main user interface

## WordPress

Extensions for WordPress are mainly used to bridge communication between CKAN and WordPress. These extensions are used to fetch and renders portal statistics, roles/categories and recommend datasets. In the alpha release, all the functionalities are encapsulated inside the theme.

## CKAN Extension

<sup>3</sup> <https://jquery.com/>

<sup>4</sup> <http://getbootstrap.com/>

New plugins were developed to support rich metadata specifications, check metadata completeness, dataset linking, personalized search, a recommendation based on profile and analysis. In most of the cases, CKAN API and Data API were used to extend CKAN capabilities. CKAN offers two kinds of API; the standard API that uses basic cataloging features and Data API<sup>5</sup>, which provides advanced query and search capabilities over structured data stored in CKAN. Data API is an additional extension and optionally requires DataPusher<sup>6</sup>. Extensions in CKAN are Python packages that extend one or more plugin interfaces. New TET metadata schema was created using existing CKAN plugin called ckanext-scheming<sup>7</sup>. Additional validators were added to ensure strict adherence to TET metadata schema.

CKAN default profile was extended by extending UserController. To support personalized search additional fields for categories and roles were added to the solar index. The search behavior was modified with IPackageController, which is a hook to package controller. For quick analysis over data, PivotTable extension was created, which is a powerful data summarization and visualization tool for tabular data. The extension uses Data API if enabled or else uses public dataproxy<sup>8</sup> instance to fetch structured data from tabular datasets. PivotTable is a view extension and implements IResourceView interface. It can be used to analyze tabular data like csv, tsv and xls. PivotTable.js library is used to render Pivot Table in the browser, additional renders for D3 and C3 are added to support for better charting and visualization of the data.

## 5.4 IMPLEMENTED FEATURES

This section lists the features that are implemented as part of TET alpha release.

### **Content management system support**

CKAN Integration with content management system enables publishers to publish content related to datasets and publish updates related to the portal in an easy way. TET WordPress plugin seamlessly integrates TET enabled CKAN and provides rich content publishing features to publishers and an elegantly organized entry point to data portal.

### **Better metadata**

More metadata fields are added to dataset upload form to enable publishers to enter rich metadata that will help users in the discovery of relevant datasets and in getting a better understanding of the datasets. The metadata fields cover the following areas and is aligned with W3C guidelines for publishing data on the web.

1. Basic details related to the dataset
2. Target audience of the dataset
3. Theme/Category
4. Versioning
5. Provenance
6. Geospatial coverage
7. Temporal coverage

---

<sup>5</sup> <http://docs.ckan.org/en/latest/maintaining/datastore.html#the-datastore-api>

<sup>6</sup> <http://docs.ckan.org/projects/datapusher/en/latest/>

<sup>7</sup> <https://github.com/ckan/ckanext-scheming>

<sup>8</sup> <https://github.com/okfn/dataproxy>

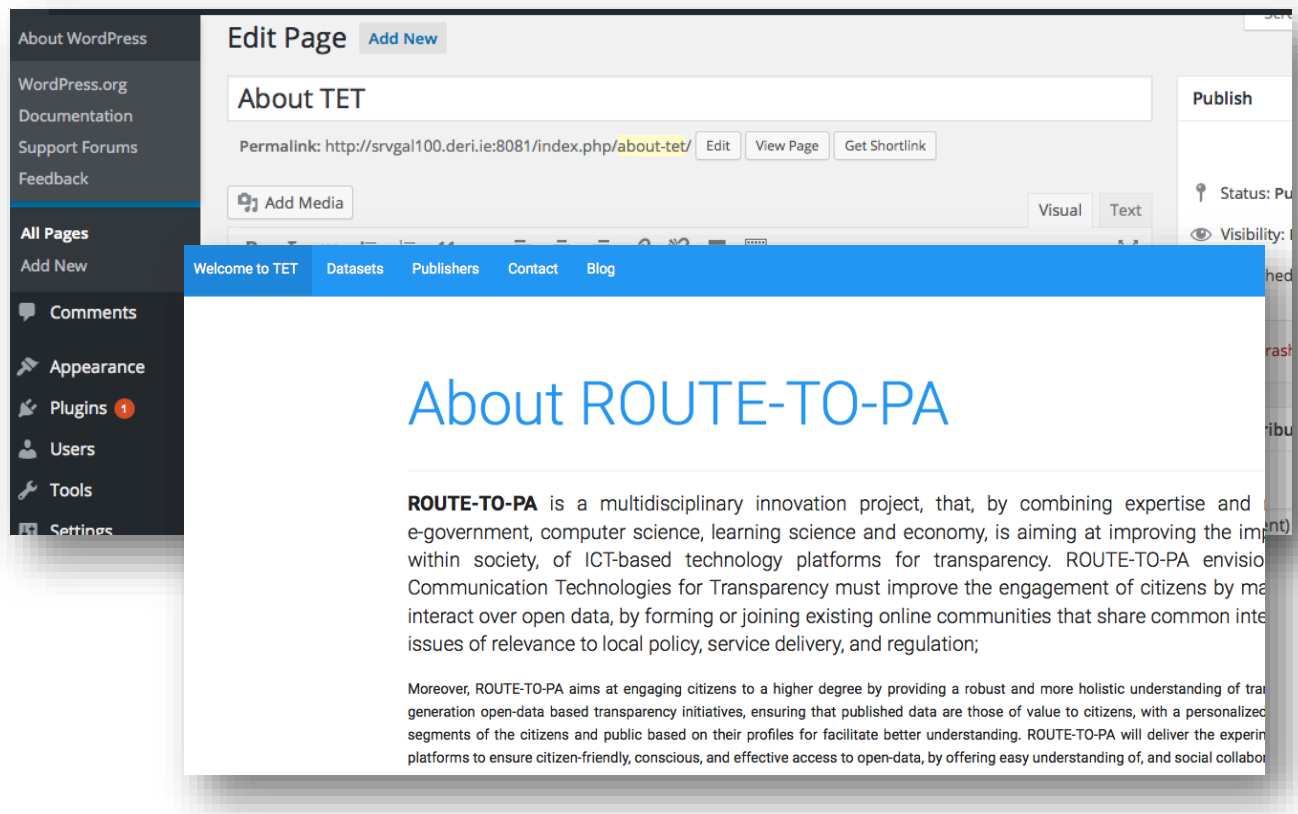


Figure 7 TET content publishing interface

### Metadata quality check/validation

Extra validations to dataset entry form are added to prevent data entry errors and to ensure consistency. Quality check indicators guide publishers about the quality of metadata being entered. The feature also helps users in assessing the metadata quality of dataset.

Figure 8 Metadata quality indicator

### Personalization for no-logged in users

The feature enables users to search based on their profile or based on the desired category. Users can select appropriate profile from the list of profiles provided or can select the category they are interested in from the list, based on the selected user; personalized results are displayed.

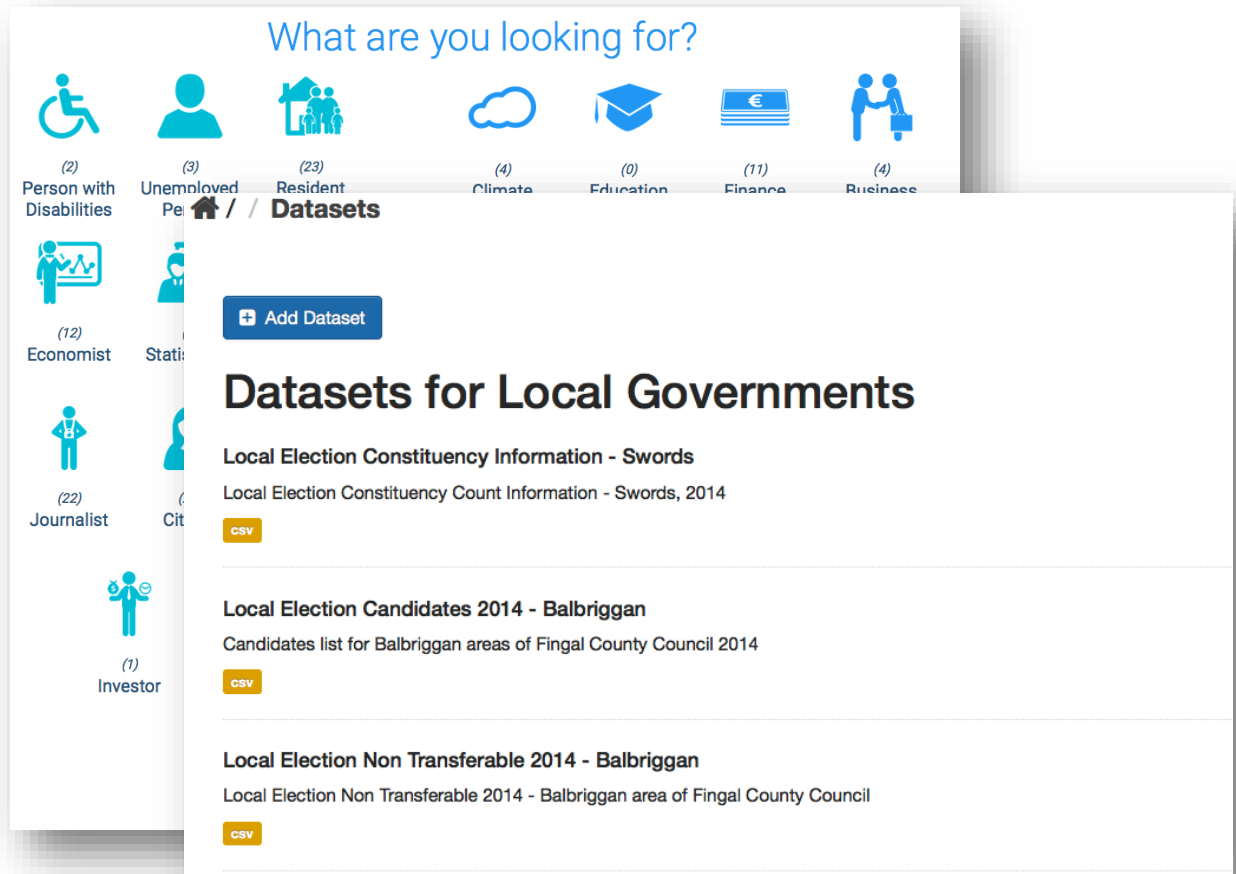


Figure 9 TET Personalization

### Datasets linking

Dataset linking feature allows users to specify explicit links between datasets, which can be exploited for recommendations and data integration purposes.



**Related Datasets:**

**\* License:**

- capital-programme
- cemeteries
- customer-service-request-fixyourstreet-ie

**\* Organization:**

Figure 10 Datasets linking

### Detailed user profiles

Default CKAN user registration page is modified to allow more details related to the user to be captured; the feature plays an essential role in creating a personalised user experience.

[Home](#) / [Users](#) / [admin](#) / **Manage**

### Change details

**\* Username:** edo

**Full name:** Edo Osagie

**Role:** Citizen

**Category of Interest:** Local Governments

**Twitter ID:** eg. RouteToPA

**\* Email:** edo.osagie@insight-centre.org

### Change password

**Old Password:**

**Password:**

**Confirm Password:**

**Delete**    \* Required field    **Regenerate API Key**    **Update Profile**

Figure 11 Detailed User Profile

## Personalized search and recommendation

The personalized search feature allows logged-in users to get personalized search based on details provided in their profile. In addition, logged-in users are provided with personalized recommendations based on their profile details.

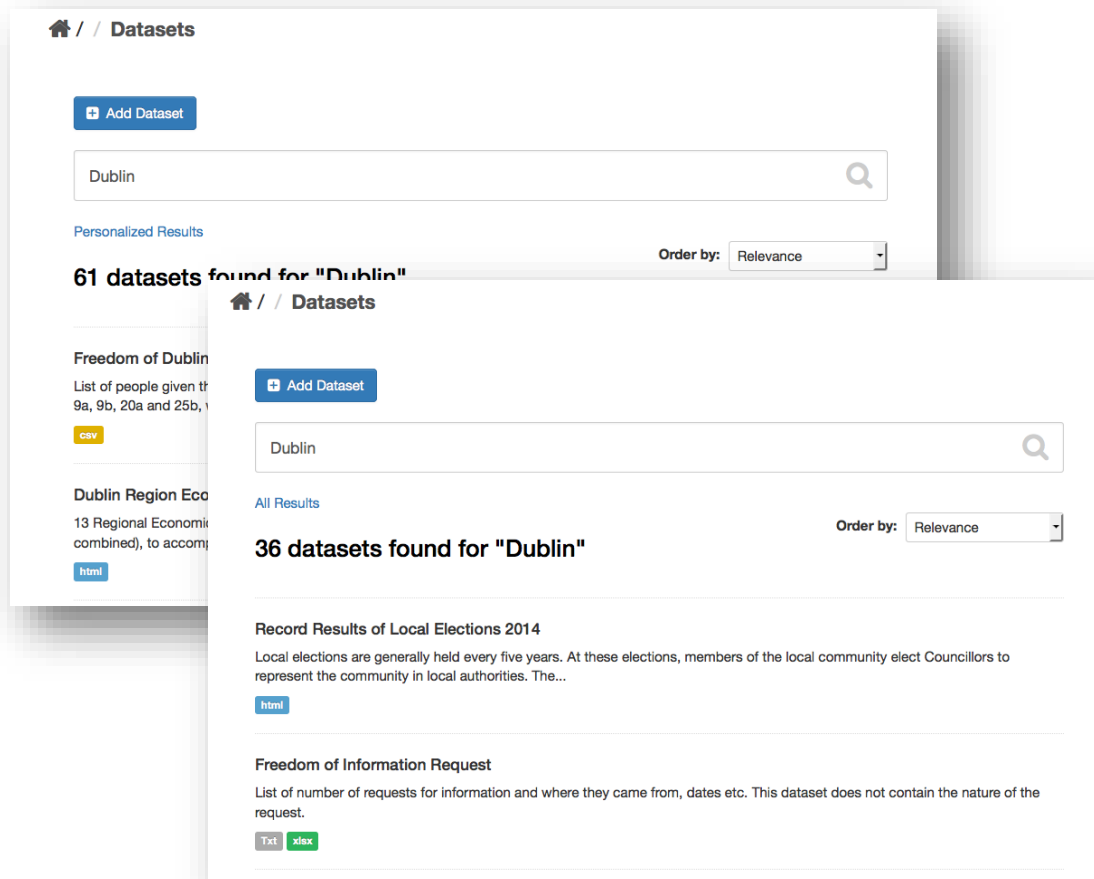


Figure 12 Personalized search and recommendation

## Recommendation for related datasets

With this feature application recommends more datasets based on the current selection and other contextual information. The feature guides users to find potentially useful and relevant datasets.

## PivotTable

CKAN platform lacks data analysis capabilities, essential for working with data. To overcome such limitation as the first step, we added the PivotTable feature. PivotTable features allow users to view, summarize and visualize data.

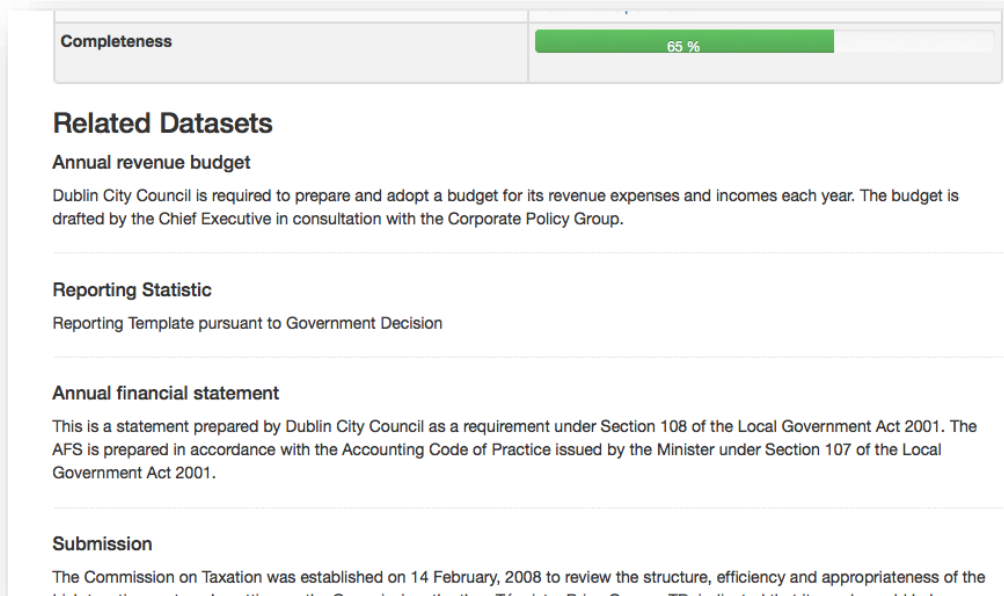


Figure 13 Recommendation for related datasets

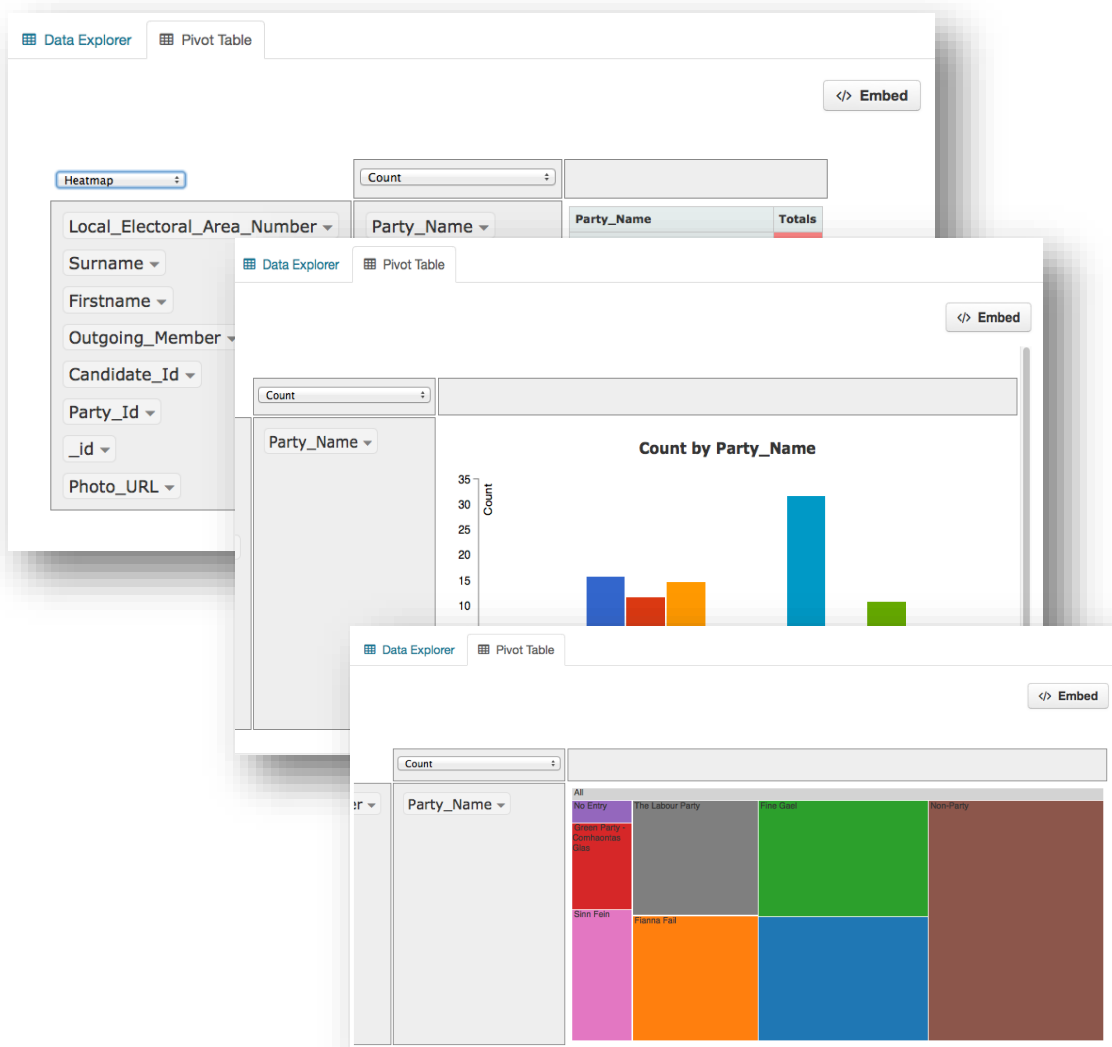


Figure 14 TET pivot view

## 5.5 INTEGRATION BETWEEN TET AND SPOD

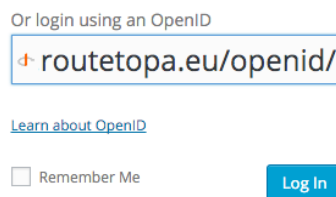
TET and SPOD are based on popular open source platforms, and their aim is to facilitate citizens' comprehension of open data, allowing user-friendly social interactions and collaboration. In order to achieve these goals in the alpha version of TET following elements were developed.

In order to enable seamless user experience across TET and SPOD, a unified user interface was designed and implemented. Both TET and SPOD have been specifically designed within the ROUTE-TO-PA project, and they are based and compliant with the material design<sup>9</sup>. In example authentication page look and feel is the same in both platforms, and the switching between pages is seamless. Moreover, at this stage, Data Stories can be created by embedding visualizations from SPOD in a blog post.



*Figure 15 SPOD visualisation embedded on TET*

OpenID<sup>10</sup> is an open standard that allows users to be [authenticated](#) by co-operating sites. SPOD OpenID authentication provider is used in order to enable single sign-on feature across ROUTE-TO-PA platform. New CKAN extension for OpenID was developed, as this feature was no longer supported in by CKAN.



*Figure 16 TET OpenID login box*

Improved visualisations, UI widgets and enhanced integration between TET and SPOD will be the focus of the next release of TET. The primary goal of the TET beta release (January 2017) will be on creating a provision for advance analytics and refinement of existing features.

<sup>9</sup> <https://design.google.com/resources/>

<sup>10</sup> <http://openid.net/>

## 6 DEPLOYMENT

---

This section defines the system requirements for TET, its deployment process and discusses the issues that may be encountered during the deployment process.

### 6.1 DEPLOYMENT REQUIREMENTS

TET implementation is platform independent and can be deployed in a variety of environments. The preferred environment for TET is Ubuntu Linux 12.04 64bit which has been used for development, deployment and testing of the TET alpha version. In addition java, virtual machine is required to be installed on the server. Other software requirements are listed in Table 4.

*Table 5: Software Requirements*

Software Requirements	More information
Web Server: <ul style="list-style-type: none"><li>• Apache 2</li><li>• PHP5.3</li><li>• Python 2.6.x/2.7x</li></ul> Indexing: <ul style="list-style-type: none"><li>• Apache Solr</li></ul> DB Server: <ul style="list-style-type: none"><li>• Postgres 9.3</li><li>• MySql Server</li></ul>	<a href="https://httpd.apache.org/docs/2.2/">https://httpd.apache.org/docs/2.2/</a> <a href="http://php.net/releases/5_3_0.php">http://php.net/releases/5_3_0.php</a> <a href="https://www.python.org/doc/">https://www.python.org/doc/</a>  <a href="http://lucene.apache.org/solr/">http://lucene.apache.org/solr/</a>  <a href="http://www.postgresql.org/docs/">http://www.postgresql.org/docs/</a> <a href="http://dev.mysql.com/doc/">http://dev.mysql.com/doc/</a>

The required capacity of the system varies and depends on the expected workload; at a minimum, a server should have 2 CPU cores, 4GB RAM and 60 GB of disk space. Table 6 provides estimates related to hardware requirements. The medium size workload should be sufficient for the city level portal while large instance should be sufficient for a national portal. Table 6 provides recommendations for hardware requirements.

*Table 6: Hardware Requirements*

Workload	Recommended hardware
Small	2 CPU cores, 4GB RAM and 80 GB of disk space
Medium	2 Servers with 4GB RAM and 80 GB of disk space
Large	2 Servers with 8GB RAM and 160 GB of disk space

For better performance and security the database and solr should be deployed on separate server,

## 6.2 DEPLOYMENT PROCESS

Deployment process involves setting up the server, installing the required software and relevant TET components and configuring the instance. TET can be either used as a fully integrated open data solution, or some components can be adopted to enhance the capability of existing portal. Detailed documentation for installation is provided with each release and can be also found the code repository. Each repository contains detailed readme file, which explains in detail the installation and configuration options available. Table 7 lists and corresponding links to the code repository of the projects that are produced as part of TET alpha release.

*Table 7: Links to the repositories*

Project Name	Project Type	Description	Link to repository
TET CKAN Extension	CKAN Extension	Extends CKAN capabilities with rich metadata specification, personalized search and recommendations.	<a href="https://gitlab.insight-centre.org/egov/ckanext-routetopa">https://gitlab.insight-centre.org/egov/ckanext-routetopa</a>
RTPA WordPress Theme	WP Theme	RTPA theme for WordPress	<a href="https://gitlab.insight-centre.org/egov/routetopa_wptheme">https://gitlab.insight-centre.org/egov/routetopa_wptheme</a>
RTPA CKAN Theme	CKAN Theme	RTPA theme for CKAN	<a href="https://gitlab.insight-centre.org/egov/ckanext-rtpa_theme">https://gitlab.insight-centre.org/egov/ckanext-rtpa_theme</a>
PivotTable Extension	CKAN extension	Provide easy analysis and visualization over tabular data using PivotTable	<a href="https://gitlab.insight-centre.org/egov/ckanext-pivottable">https://gitlab.insight-centre.org/egov/ckanext-pivottable</a>
CKAN OpenID Extension	CKAN Extension	Provides authentication with OpenID	<a href="https://gitlab.insight-centre.org/egov/ckanext-openid">https://gitlab.insight-centre.org/egov/ckanext-openid</a>
RTPA-devenv	Vagrant Project	Setups the development environment for TET	<a href="https://gitlab.insight-centre.org/egov/rtpa-devenv">https://gitlab.insight-centre.org/egov/rtpa-devenv</a>

TET development environment is managed with Vagrant<sup>11</sup>. RTPA-devenv contains the Vagrant configuration and associated data file, which are required to configure the development environment. Once the development is finished, the application will move to a staging area where new features are tested before moving to the production environment. The process is illustrated in Figure 7 TET deployment process.

---

<sup>11</sup> <https://www.vagrantup.com/>

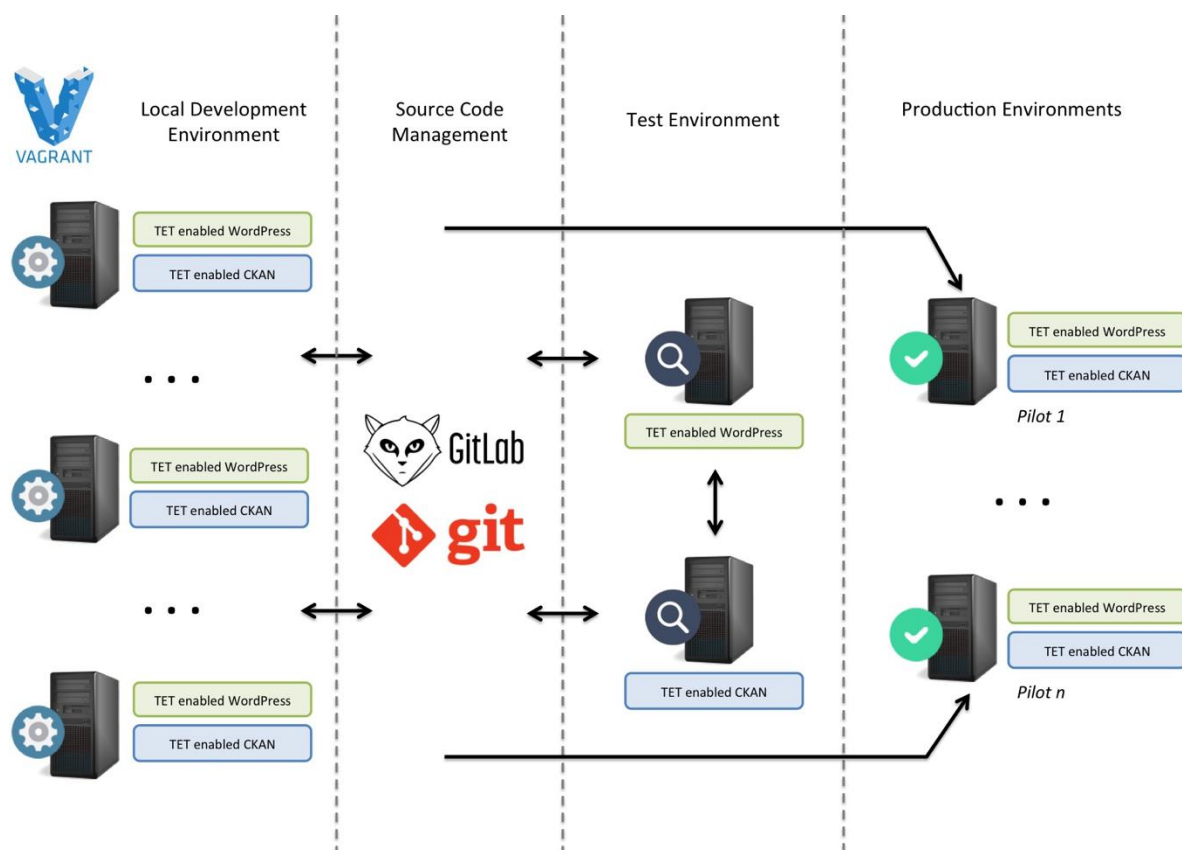


Figure 7 TET deployment process

### 6.3 DEPLOYMENT ISSUES

Migration, software versions mismatch and metadata schema mismatch are common issues that can be encountered during the deployment process. Data migration from existing platform could be tricky and depends on the source platform. Most of the open data platforms have implemented a harvesting feature that enables transfer of the data from one platform to another. In some cases, custom scripts may be required in order to transfer the data.

TET has been built on CKAN in version 2.4.0, and some features may not be compatible with its previous version of CKAN. Therefore, in some cases, CKAN version needs to be upgraded in order to maintain compatibility with developed extensions.

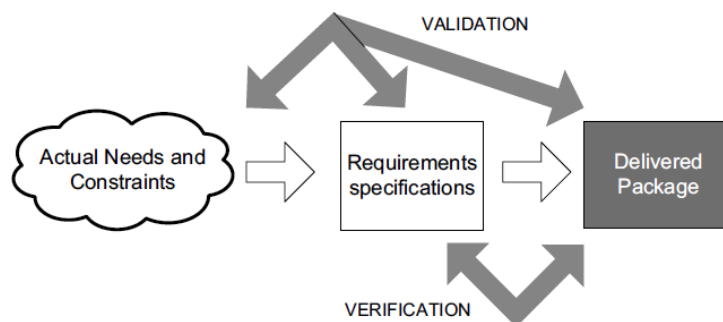
Each organization has their own way of describing metadata associated with the data. TET defines a metadata schema that is compliant with the W3C recommendations for data on the web. Changes may be required in order to adopt the TET schema ideally. However, it is also possible to configure the TET schema in such a way, that it could be easily adopted with existing metadata schema.

Categories and roles are essential metadata fields for enabling personalized search and recommendations. These fields could be missing, or list of categories and roles may be different in the existing system. While the categories issue can be easily fixed by changing the settings in the configuration file, the roles issue would either require manual data entry or custom scripts could be used to add these details.

## 7 VERIFICATION AND VALIDATION

---

Software product development generally combines two groups of tasks – the **Verification** processes and the **Validation** processes. The verification processes describe the activities of consultation with users, gathering their requirements and constraints; and checking and analysing the requirement documentation to ascertain alliance of a system with users' expectation, usability and performance. In contrast, validation processes refer to the activities of checking and testing the functionalities to ascertain alliance and correspondence of the system (product) with its specifications derived during verification phase (Tutorial points, online). The figure below (*Figure 17*) shows a two-leg transition processes involved in the conversion of user needs into the requirement specifications which are needed to build a software package as a product that delivers the needs of the client. It tends to separate verification processes from the validation processes. It shows that verification processes occur in the second leg of the journey between requirements specifications (inclusive) and the final product release while the validation processes span across the entire journey.



*Figure 17: Validation and Verification Processes in Software Development (Baresi, L & Pezze, M. 2006)*

The essential point of interest for the ROUTE-TO-PA teams in the V-Model is the fact that the model clearly defines the various tasks involved in each segment of the framework in line with our choice. The V-Model defines the tasks of Verification as including the following tasks according to the (Tutorial points). Some of these tasks such as D2.1, D2.2, D2.3 and D2.4. have been delivered in previous deliverables of ROUTE-TO-PA.

### 7.1 SYSTEM VERIFICATION

In software engineering, the commonest definition of Verification is the answer to the question: **“Are we building the system right?”** (Easterbrook Steve (2010)). Verification is the process of carrying out research, examination, etc., required to prove the authenticity or validity. It seeks to obtain evidence that confirms the accuracy or truth of something. The tasks briefly described below are common tasks that are carried out in verification processes of ROUTE-TO-PA.

- **Business Requirement Analysis:** Consult with users which enhance our ability to capture their needs and constraints. These needs and constraints were documented, analysed, produced requirement specification from it and then communicated to users and other stakeholders for full understanding and agreement. Because users do not and cannot understand all possibilities of the design



of the product or additional functionalities, the product creation team added other complementing requirements to augment and strengthen the users' perspective of the requirements. The other task required to be done at this stage was the preparation of both the test plan and the User Acceptance Test (UAT) Plan. The creation of a test plan at this stage is a good practice because the understanding of the requirement specification is a necessary input in the creation of proper test cases from the use cases that were produced during this phase.

- **System and Architectural Designs:** The design phase began once a good understanding of the requirement specifications was established from the requirement analysis. It involves the evaluation of the needed hardware and other applications for the complete setup of the product needed by the users. In the Architectural design phases of TET, System design has been broken down into modules to focus on the details of the different functionalities. The High-Level Design (HLD) stage also has been reviewed and designs modalities to integrate the internal and external systems.
- **Module Design:** This is the low-level design stage in which smallest units of the systems – the component parts by functionalities are designed and tested. The tests that are done at this level are low-level testing such as unit testing; however, other tests include integration testing and automated testing. Testing during module design stage provides the advantage of eliminating the majority of the design errors.

So far, in the course of ROUTE-TO-PA product development, several deliverables up to deliverable D2.4 (Requirement Specification and Use Case Models) focused on the first leg of the journey from user needs and constraints to a final product designed to satisfy the needs. Through this same deliverable (D2.4) we have entered into the second leg that involves coding (verification) and verification of codes towards final package delivery in an approach in which both the validation and verification processes will complement each other.

## 7.2 TESTING FRAMEWORK

Having developed the requirements specification documentation for ROUTE-TO-PA in (D2.4) and have communicated it to all stakeholders and together agreed on it, we decided on building the system and testing each build in an iterative manner. First, we considered building a plan – the Test Plan that gathers input from the methodologies of code development, and testing approaches that we have chosen to use in the project implementation phase (Implementation).

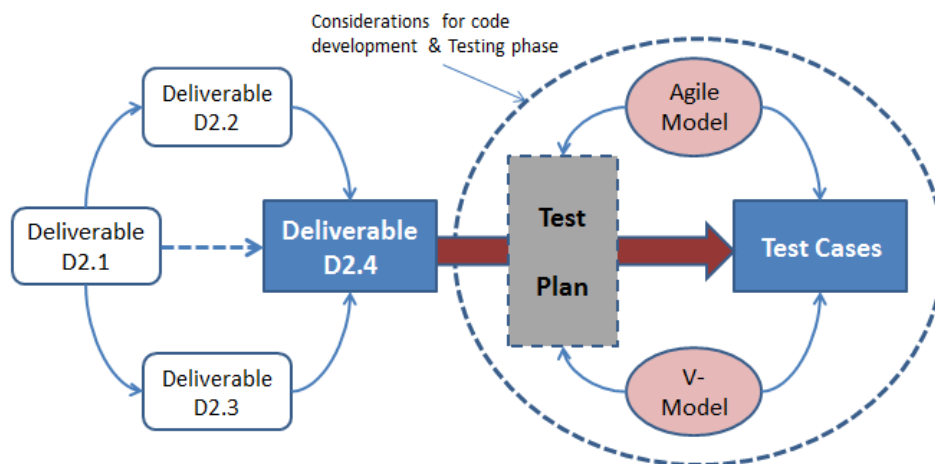


Figure 18: Considerations for Test Plan, Code Development & Testing

In developing our Test Plan, we considered inputs from the Agile methodology and the V-Model that we have chosen to use in the project, and these were also relevant factors in creating the test cases. Due to the choice of the agile approach we created test cases that are independently testable functionalities based on the use cases. These test cases are presented in the appendices section of this document.

## V-MODEL TESTING

The V-Model, also known as the Verification and Validation (V&V) model is an SDLC approach whereby execution of software development processes occurs in a V-shaped manner (Tutorialpoint). The essential note of this model is the simultaneous execution of the testing phase along with the corresponding development phase. The code development and the test planning and testing (verification) carried out at the same time – that is, the test planning/design and the testing task are directly associated with every coding task. To enhance the possibility of executing testing activities with the coding activities, requirement analysis should produce requirement specifications that identify *independently testable features* of the software product (Baresi & Pezze, 2006).

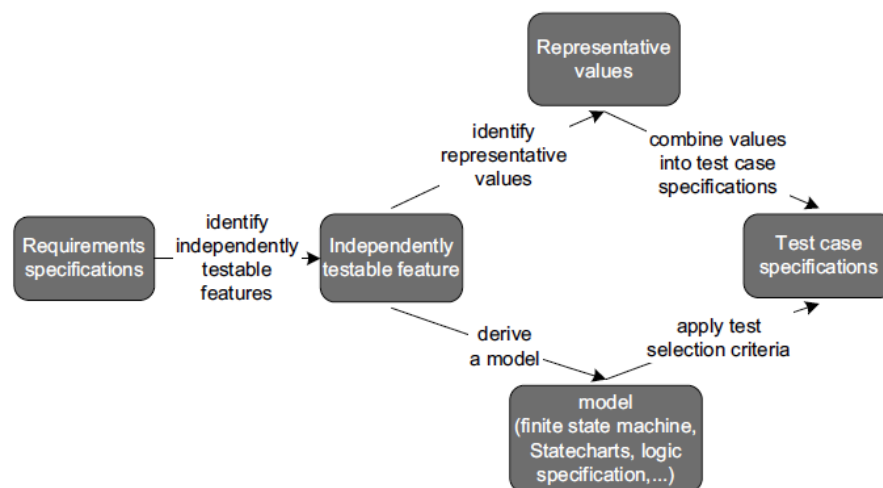


Figure 19: A systematic approach to system testing. Source: (Baresi & Pezze, 2006)

The testable features are made up of representative values [the desirable functional values] that can be redrafted into the test cases. The features are also adapted to be model-driven and which are subjected to test criteria that suit the test case specifications. This guide enables software development team to improve their chances of successfully applying the V-model. In ROUTE-TO-PA project, we have carefully carried out the requirement analysis and produced the low-level specifications that are based on the features or functional units of the products under development. The use cases were designed one after the other based on the user requirement augmented with the system qualities to facilitate User Acceptance.

The two sides of the V-Model represents different streams of software development activities – while the left side talks about the Verification activities, the right side is about the validation activities. Figure 20 shows the diagram of the traditional or basic V-model. In this diagram, each verification activity having a corresponding testing or Validation activity as well as the necessary planning task for the testing processes.



### 7.3 TEST APPROACH (AGILE METHODOLOGY)

The Agile Model is the most applicable and useful model from our review of several models, and we have planned to adopt it in the development of ROUTE-TO-PA system.

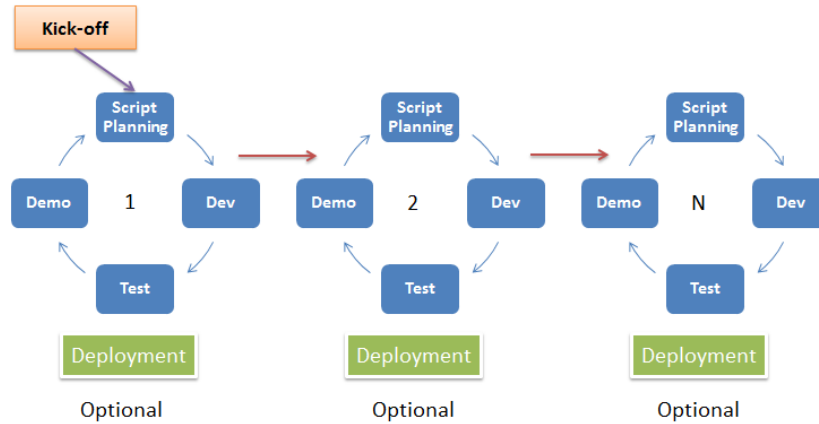


Figure 21: Agile Model - Adopted from ISTQB examcertification.com

As seen on Figure 21, system implementation cycle (1) starts with script planning both the codes and the test scripts. Next is the coding (or development) phase of the system feature or functionality based on each use case. After that, we put the developed feature through a test to verify the code correctness or validate error-free functionality using the test scripts. If any bug is found during the test, the cycle (2) is repeated to correct the bug. This cycle will continue to 'n' until all bugs are corrected and the codes for the feature are verified in alliance with the requirement specification and functionality is validated in accordance with the user needs.

We plan to begin the agile development cycles from the very low-level features and test them as they are developed build-by-build so that any bugs found can be dealt with much more easily.

#### **Advantages and disadvantages of Agile Methodology**

By adopting the Agile method of software development in ROUTE-TO-PA project, the team hopes to explore the advantages offered by the approach to improve the quality of the product and reduce budget time and cost. According to The International Software Testing Qualifications Board (ISTQB) several advantages (and disadvantages too) accrue from the adoption of the Agile approach and these advantages include

1. The rapid and continuous delivery of working software product produces customer satisfaction.
2. Interaction of all parties (stakeholders: Customers, developers and testers) concerned is important and not just a focus on processes and tools. The interaction produces value as the process progresses. There is improved communication due to frequent face-to-face conversation and communication.
3. Early delivery of working software is guaranteed, and this keeps the customer or user happy.
4. The requirement specifications for software development are unavoidably changing as development progress. The agile method offers an opportunity for the team to adapt to the changing circumstances enhanced by the short feedback loop.

On the disadvantages side, firstly, Agile method may not prove effective in cases of large software development project because it is difficult to assess the effort required at the beginning of the development cycle. Secondly, design and documentation may not get enough emphasis as things move very rapidly. The consequence of this

situation can also result is that the project may go off-track. In ROUTE-TO-PA project, we have made an effort to avoid these shortcomings. We have developed detailed documentation to serve as guides for developing the required product and to meet quality according to user demands as reported in deliverables:

- D2.1 – State-of-the-art Report and Evaluation of Existing Open Data Platforms
- D2.2 – Analytical Framework and Initial Scenarios on Open Data and Transparency
- D2.3 – User stories on Open Data and Transparency, and
- D2.4 – Requirement Specification and Use Case Models
- Others include ROUTE-TO-PA Test Plan and ROUTE-TO-PA System Development and Testing Framework (this report) containing the Test Cases.

The development and application of the aforementioned reports and, in particular, the Test Plan, are effective strategies that can ameliorate the impacts of the difficulty in assessing the effort required in Agile Approach (section 7.4). Due to the fact that ROUTE-TO-PA is a deliverable for users (the ordinary citizens) considered to be non-tech savvy, changes to the initial specification documentation are inevitable. Agile method will enable the freedom to implement the new changes without incurring too much cost, as a result of, frequent feedbacks and short delivery loops

## 7.4 TEST PLAN

This section presents the plan for testing in the development phase of the project and is designed to follow the template created by the Institute of Electrical and Electronics Engineers (IEEE 829). The template enables us to identify and clearly set out the various aspects of the testing procedure prior to the commencement of the testing phase of this project. The Test Plan development task is based on the understanding that the testing procedure would go side-by-side with the code development (implementation) phase following the popular V-model of software testing procedure. In addition, to the above, test planning is a strategy that introduces some advantages into the testing procedure and thus, testing is not just exploratory in nature – i.e. a tester simply tests a product in the normal sense of it. Exploratory testing is therefore not stringent enough to ensure a proper, complete testing cycle according to Brissonnette (2015) who proposed the following advantages of using test plans for software testing.

**Assurance of complete testing** – Test plan helps to ensure that project team has tested all aspect of the product and that not data flow or use case is omitted through a build-by-build testing approach in each phase of the development process.

**Enhancement of test structure and organisation** – Test plan enhances the organisation of the testing process and thus allow better test flow as the implementation progresses. It is a perfect way to organise a test and enables testers to report all of the bugs they find randomly into a bug tracking record.

**To prevent “Tunnel Vision”** – Tunnel Vision refers to a defective eye sight whereby an object not close to the field of view cannot be seen properly or clearly. In the informal usage, tunnel vision usually occurs in a monotonous exercise in which the tendency to focus exclusively on a single or on a limited objective, view or possibility rather than the many instances exist. In this sense, a tester may experience “Tunnel Vision” if the tester has spent too much time on a certain project that he/she becomes too comfortable with it to the

extent that the risk of missing out a system bugs is increased. A common example of this is that if a tester has been used to starting a test from the beginning, at the time a new but minor functionality is introduced or removed, the tester may not spot it unless the change is significantly obvious. With a test plan, Tunnel Vision can be eliminated.

**Test Plan as a guide for Testing** – Test plan serves as a reference guide for testers in times of challenges, and when they fall into trouble thus it serves as a legal document of protection between developer and tester in time of differences

**Facilitation of project management** – A Test Plan helps a project manager to understand the scope of tests job and items to be tested as well as those not to be tested, and to manage the scope and flow of testing exercises in alignment with development jobs. It also enables project managers to monitor budget properly in accordance with the overall project objectives and deliverables.

As a summary exerts from ISTQB, the test plan is created first to guide our thinking and conform to the standardised testing procedure as a strategy against challenges in the process. Second the test planning task and the plan itself both serve as means of communication within the project team involving the developers, testers, and other stakeholders. Finally, the test plan is a document that helps in the management of change during the testing phase of the software development lifecycle. The purposes of adopting software test plan in this project are in accordance with the recommendations provided by the Softwaretestingtimes.com

- To ensure all Functional and Design Requirements are implemented in accordance with the specification of the existing project the documentation especially the use case documentation deliverable D2.4 (Deliverable 2.4 Requirement Specification and Use Case Models)
- To provide a procedure or guide for Unit build-by-build progressively and System Testing at the end of the development phase.
- To identify the documentation process, scope and monitoring for Unit and System Testing.
- To identify the suitable testing methods for Unit and System Testing from the outset.

Also, this section details the breadth and length of the testing tasks that will be performed in ROUTE-TO-PA project. The project approach is Agile and in addition to this framework, the team will adopt the V-Model for application in the testing phase of the project. The requirement specification document – deliverable D2.4 (Requirement Specification and Use Case Models for TET and SPOD Subsystems) provides the bases for the scope determination as it also provides the ROUTE-TO-PA application functionalities to be tested. The less sophisticated single component functionalities requiring little interactions with other components of the system will be the ones that will be developed at the very beginning of the implementation phase. This stage of the implementation phase will be followed by other stages implementing the functionalities, components or subsystems in accordance with their complexities or importance of the linkage role they may play in the software system as a component. The iteration strategy thus utilises the *build-on-previous* version agile methodology to develop the product in an incremental manner from start to finish of the implementation process. Likewise, testing will follow the cycles of iteration by testing every build of the components.

In specific term, testing objectives require that verification and validation processes will be carried out in this project to cover various categories of ROUTE-TO-PA web-based application as a product. As much as possible, the team hopes to build the ROUTE-TO-PA components by following the *clustering* of the requirement specifications earlier produced in deliverable D2.4 (Requirement Specifications and Use Case Models). However,

due to continued requirement management and communication activities with subsequent feedbacks from stakeholders, the earlier clustering has undergone considerable modification and fine-tuning to reflect better understanding and general agreement by all stakeholders. The result of the modifications which will better enhance the implementation phase is shown in *Table 8*.

#### 7.4.1 TEST ITEMS

In ROUTE-TO-PA, the components or items that will be tested include but not limited to those in *Table 8*, but a subset of these components are in Alpha version ready for testing and are shown in Table 1 above. As the need may arise from iteration processes, requirements may be reviewed to redefine them, remove or add some just to ensure that user needs and other stakeholders need are validated with the solution (product) requirements.

*Table 8: Components of Development and Testing (subject to changes if necessary).*

Test Case ID	Ref Use Case ID	Use Cases (to be tested) / (Items to be tested)	System (Interface) (Sub-sys)	Ref User Story ID
TC14-UC14	UC14	Link a Dataset	TET	S8.1
TC1-UC1	UC1	Check Metadata Completeness	TET	S29.1
TC23-UC23	UC23	Login	TET	S28.1
TC2-UC2	UC2	Add/Update/View Provenance and metadata to related to a Data Set	TET	S29.3
TC3-UC3	UC3	Check Dataset Quality	TET	S29.2
TC17-UC17	UC17	Personalize Search	TET	S28.2
TC22-UC22	UC22	Post on SPOD	SPOD	S28.3
TC24-UC24	UC24	Enrich Profile	TET	S38.1
TC8-UC8	UC8	Using External Social Media for Login Authentication: e.g. Facebook (External Social Media Authentication)		S3.0
TC20-UC20	UC20	View a Chart	TET	S2.2
TC21-UC21	UC21	View Infographic	TET	S2.2
TC11-UC11	UC11	Query Data	TET	S32.2
TC18-UC18	UC18	Receive Notification for New Datasets		S3.1
TC19-UC19	UC19	Request Recommendations for Data Sets	TET	S26.5

TC26-UC26	UC26	Add an Event	SPOD	
TC27-UC27	UC27	Create a New Topic Enabling an Anonymous Posting	SPOD	
TC4-UC4(A)	UC4(A)	Provide Supporting Content on SPOD Post	TET / SPOD	S2.3
TC5-UC5	UC5	Provide Supporting Content on SPOD Post (Share Link to Dataset/File via SPOD Post)		S35.1
TC10-UC10	UC10	Analyse a Dataset	TET	S1.4
TC12-UC12	UC12	Integrate a Dataset	TET	S8.1
TC15-UC15	UC15	Add a Dataset to Compare	TET	S13.9
TC16-UC16	UC16	Compare Datasets		S13.9
TC28-UC28	UC28	Digest Information in a Private Space		
TC29-UC29	UC29	Enable Rewarding Content on SPOD		
TC4-UC4(B)	UC4(B)	Share a dataset with a visualisation on SPOD	TET/SPOD	S2.1
TC6-UC6	UC6	Take a Screenshot and Annotate it		S7.1
TC7-UC7	UC7	Attach Widgets and Tags Tools on SPOD Forum		S31.1
TC9-UC9	UC9	Voting on SPOD FORUM		S2.6

The above component of the project has been aligned with the objectives of the programme regarding the transparency enhancement and social engagement over open data. These have in turn been tied back to the Societal values of Open Data as demonstrated by the three models of democracy namely: Monitorial Democracy, Deliberative Democracy and Participatory Democracy.

#### 7.4.2 FEATURES NOT TO BE TESTED

ROUTE-TO-PA component verification and validation from the point of view of our involvement as a partner in the project will be focused on the Transparency Enhancing Toolsets testing. Some TET functionalities have connections with the SPOD interface in an attempt to satisfy user needs, for example, using SPOD to send the TET-analysed file. These cases normally tag as TET/SPOD activities are designed to test TET/SPOD seamless integration capabilities to execute a specific role. While TET-specific functionalities and capabilities will be covered in this place by the Insight (NUI Galway) team, the SPOD specific functionalities of ROUTE-TO-PA solution will be covered by the Prato Team. Although Prato will focus on the testing of SPOD specific features, however, the team will collaborate with Insight in testing TET/SPOD features.



#### 7.4.3 TEST TOOL SELECTION

Since ROUTE TO PA project is using an agile approach, with weekly iterations. At the end of each week, the requirements identified for that iteration will be delivered to the team and will be tested. Exploratory testing will be employed a great deal for the testing components. Tests for planned functionality will be created and added to the test documentation and will be updated when there is an iteration of the system. Although testing activities begin with the manual approach, however, the automated testing method will be used towards the end of the project.

**Manual testing:** This is the process of manually testing software for defects. This process requires a tester to play the role of an end user and use most of all features of the application to ensure correct behaviour. The required test tools for this manual testing exercise for ROUTE TO PA includes:

- MS Excel will be used to create test cases, log issues and document analysis of risks.
- MS Word will be used for documentation of the test plan, test checklist and summary reports.
- Skype and Email will serve as a communication tools to get feedback from the test team and other test other users involved in the testing exercise.

**Automation testing: Automated testing** can be described as an act of converting manual test cases into automated scripts that can be executed autonomously and can be conducted on a stable and non-changing application (Day, 2014). By automated, it means that the tests themselves are code. The tests can then be run over and over again with a very little effort, at any time, and by anyone.

When tests have to be run manually (with someone sitting at the computer typing the input on the keyboard), the execution of the manual tests and the examination of the results can be error-prone and time-consuming. Automation testing will be useful for ROUTE TO PA because when schedule pressures rise, manual testing often gets forgotten. Furthermore, automating tests can be very beneficial to this project due it's emphasized in agile development. Although Automating tests can be time consuming and expensive because writing an automated test can take several orders of magnitude more time (2X – 10X more) than executing the test by hand once ("Agile / Automated Testing," 2003).

The required testing tool for the automation testing exercise for ROUTE TO PA is the selenium which is an open source automation testing tool that has been agreed upon by the testing team. **Selenium** is composed of multiple software automation tools such as selenium IDE, selenium RC (selenium 1.0) and selenium webdriver(selenium 2.0).

- Selenium IDE is an integrated development environment to build the test scripts. It is a firefox plug-in which allows the record, edit and debugs the selenium test cases. It is responsible for all the recording of all performed actions by the end user and generates the test scripts (Gojare, Joshi, & Gaigaware, 2015).
- Selenium remote control (RC) was regarded as the main selenium project for a long time due to it slow features; it requires to start the server before executing the test scripts. It doesn't support the Ajax applications. To avoid the limitations of selenium RC, selenium webdriver was invented by merging together of selenium and webdriver (Gojare, Joshi, & Gaigaware, 2015).
- Selenium webdriver also known as selenium 2.0 directly communicate with the browser, so it is faster than selenium RC. It also supports multiple web browser and Ajax applications. The main goal of the selenium webdriver is to improve the support for modern web application testing problems. Furthermore, it also supports the multiple languages available for writing of test scripts which makes the selenium webdriver's API simpler to selenium RC (Gojare, Joshi, & Gaigaware, 2015).

After a brainstorming session within the testing team where comparison, pros and cons of the three types of selenium were considered, it was agreed upon to use the new version of the automation testing tool known

as selenium webdriver 2.0 mainly because of the new added features and the support it for multiple programming language and web browser applications.

**Bug tools:** Since the objective of the testing activities is to identify any error or defects on the TET platform. Therefore, any bug or defect found during the testing phase or any issue raised will be recorded on the following excel sheets created:

- Defect log
- Incident report log

#### 7.4.4 TEST DATA

Test data is a data employed during the testing phase of the system. For testing to be conducted the test data must be provided. Test data can be categorised into two namely valid data and invalid data.

- Valid data can be described as the entry of acceptable data formats into the system. Its test cases are executed to confirm if the system accepts the valid data values.
- Invalid data, on the other hand, can be described as the entry of data in the wrong format into the system. Its test cases are executed to confirm if the system rejects invalid data values.

*Table 9: Test Data for Alpha Version of TET Components*

Dataset types	Sources	Elements	Series	Formats
Budgets, spend, procurement	Dublinked	Purchase order	Quarterly	Csv & Excel
Budgets, spend, procurement	Dublinked	Annual financial statement	Annually	Pdf
Budgets, spend, procurement	Dublinked	Audit reports and manager's responses	Annually	Pdf
Budgets, spend, procurement	Dublinked	Report statistics	Quarterly	Pdf/ xls
Budgets, spend, procurement	Dublinked	EU IMF reporting	Annually	Excel
Budgets, spend, procurement	Dublinked	Annual revenue budgets	Bi-annual	Pdf
Budgets, spend, procurement	Dublinked	Capital programme	As required	Pdf
Council meetings	Dublinked	Meeting minutes and agenda	Monthly	Pdf
Council meetings	Dublinked	Arts, culture, recreation and community SPC	Annually/ monthly	Pdf
Local election results	Dublinked	Result of 2014 local election	Every five years	Pdf
Local community facilities, services, transport	Dublinked	Community API	Released 17/01/2011	Links
Economic monitor	Dublinked	Economic indicators	Quarterly	Pdf/links

data				
Sustainable development indicators	Dublinked	Indicators	Annually	MS Word
Customer service request	Dublinked	Public API	Daily update	API
Freedom of information request	Dublinked	FOI	Sample set released on 28/03/2013	Csv & Excel
Polling places	Dublinked	Polling stations	As required	Csv
Freedom of Dublin city	Dublinked	List of honorary awardees	As required	Csv
Lord mayors	Dublinked	List of previous lord mayors and deputy lord mayors	As required	Csv
Bye laws	Dublinked	Records of secondary laws	As required	Csv
Councillors, allowance and expenses, conferences, payments, webcasting view	Fingal open data	Records of locl elections candidates, allowances and expenses	As required	Csv

#### 7.4.5 TEST DELIVERABLES

*Table 10: Test Deliverables*

<b>Deliverable</b>	<b>For</b>	<b>Date/ Milestone</b>
Test Plan document	Test team, developers and unit leader	
Test Design Specifications (test cases)	Test team and unit leader	
Defect Logs, Incident Logs And Execution Logs	Test team	
Test Summary Report	Unit leader and developers	

#### 7.4.6 ENVIRONMENT NEED

This section provides details on the test resources required for the testing phase of TET subsystem, which includes hardware, software, network and documentation.

##### **Hardware**

The required devices accepted for the testing of the TET platform must be in good condition with a reliable internet connection, it will also support at least one of the following operating system: Android, IOS, Mac OS X and Windows. The test team should confirm that these devices should be in place for effective testing of the platform.

##### **Software**

The TET platform has been developed to work best in the latest desktop and mobile browsers, which means older browser might display differently styled, though fully functional, renderings of certain components. Below is a matrix, which shows the operating systems and the browser they support.

*Table 11: Test Deliverables*

Operating system	Chrome	Firefox	Internet Explorer	Opera	Safari
Android	Supported	Supported	N/A	Not supported	N/A
IOS	Supported	N/A	N/A	Not supported	Supported
Mac OS X	Supported	Supported	N/A	Supported	Supported
Windows	Supported	Supported	Supported	Supported	Not supported

However, TET platform should look and behave well enough in Chromium and Chrome for Linux, Firefox for Linux, and internet Explorer 7, although they are not officially supported.

#### 7.4.7 DOCUMENTATION

After each test is performed, a test summary report sheet will be completed, where the result/outcome of the test will be recorded. This summary report sheet has been designed to record information such as:

- Test stage which identifies the type of test e.g. unit, functionality, integration, performance, system regression, etc.
- Tester details, test log ID, test case ID, summary review date.
- Test summary
- Variance and assessment
- Summary of results, evaluation and corrective action plan
- Approvals

#### 7.4.8 ITEM PASS / FAIL CRITERIA

##### 7.4.8.1 TEST PHASE ENTRY AND EXIT CRITERIA

**Entry:** Before the commencement of the testing phase/activities in ROUTE-TO-PA, there are some various factors required for consideration by the test team. Some of these factors includes:

- All documentation on the requirements of the platform has been gathered
- The test team needs to be walked through the application and possess a good understanding of how the application works.
- All necessary test cases have been developed, reviewed and are ready for execution.
- The environment (web url) that will be used is configured and ready for use.

**Exit:** The exit criteria can be described as a process whereby the test team is in readiness to bring the testing activities to an end after a successful completion. To conclude test activities on TET platform, the test team must be sure of the following:

- All developed test cases were executed successfully unless all parties involved agreed to exclude some of them.

- All raised issues occurred during the testing activities were resolved and retesting were conducted to ensure that there are no pending issues.
- The testing summary report has been completed.
- All high priority defects have been fixed and closed.
- For Medium and low priority defects that have not been fixed, their impacts have been analysed, and mitigating measures are in place.

The sign off and exit criteria will come to pass after the above criteria have been satisfactorily fulfilled.

#### **7.4.8.2    UNIT TEST PHASE ENTRY AND EXIT CRITERIA**

**Entry:** Each of the functionalities of the ROUTE-TO-PA application will be tested individually by the developers according to the test case developed for each to verify that all features work as expected. The expectation of each function is to return the proper values when been asked to perform its actions. If it fails, a bug ticket will be issued and recorded in a defect log sheet, advising to be fixed and will be retested regressively.

**Exit:** To exit this phase of testing, the test team must be assured by the developer with empirical evidence that all various functionalities in the systems have been tested individually through the execution of test cases developed for each.

#### **7.4.9    INTEGRATION TEST PHASE ENTRY AND EXIT CRITERIA**

**Entry:** To verify if the different functionalities of the TET system have been integrated, the system will be tested rigorously from different angles by testers under the coordination of both the test manager and developer manager. This will ensure that the design objectives are met and that the software complies with the operational requirements. The bottom –up approach will be used during this phase which means lowest level units will be tested and integrated first before the high –level units are integrated. If an error is detected, a bug ticket will be raised and documented in a defect log sheet and advising for a fix. A regression testing will then be carried out to verify if the fix has not caused any damage to other components.

**Exit:** To exit this phase, the test team must be assured that all the units in the system are fully integrated, and the design objectives of the system have been met.

#### **7.4.10    ACCEPTANCE TEST PHASE ENTRY AND EXIT CRITERIA**

**Entry:** The aim of this phase will be to establish confidence in TET, validate the fitness-for-use of the platform and also to validate the whether the system meets the requirements for operation. To approach the activities required for this phase, people outside the developing and testing teams and other stakeholders will be invited to carry out testing activities. If an error is detected, a bug ticket will be raised and documented in a defect log sheet and advising for a fix. A regression testing will be carried out to verify if the fix has not caused any damage to other components.

**Exit:** To exit this phase, the test team must be assured that invited participants to carry out acceptance test are satisfied and agreed that the system meets the requirements of the system and also fit for use.

#### 7.4.11 SUSPENSION AND RESUMPTION CRITERIA

The test team will suspend the test activities if:

- Code freeze is not maintained: this means during testing, the developing team are not allowed to fix any bug in the application. This is because it can interfere with the testing activities being carried out and the testing team might not be able to get the clear results out of the testing procedure.
- The application crashes on a particular platform.
- The test environment completely fails.

To resume the test activities, all the raised issues that led to the suspension must be resolved, and also measures must be put in place to mitigate any unforeseen issues that might arise after the resumption of the testing activities.

#### 7.4.12 REMAINING TEST TASKS

Table 12: Other Testing Tasks

No	Task	Assigned to	Status
1	Create Acceptance Test Plan	Samuel Adebayo	In progress
2	Create System/Integration Test Plan	Ed Osagie	In progress
3	Create Test Summary Report	Ed Osagie / Samuel Adebayo	In progress
4	Create User Guide	Ed Osagie/ Samuel Adebayo	Completed

#### 7.4.13 PERSONAL AND TRAINING NEEDS

- Selection – internal and external testers: It is important to ensure that those that will be selected to participate in the testing will add value to the process. Members of the unit test team will take part in the test while external will be from other units within the insight centre, NUIG. All selected member will be allocated to different roles but with same focus of finding defects in the system.
- Training: For successful activities of testing to be carried out, some important areas needs to be addressed in terms of training. These areas are:

The developers and testers will need to be trained on how to use the TET web application and prior to final acceptance of the project stakeholders will also require to be trained on how to interact with the TET web application.

Other potential users will require training on how to use the system.

At least one developer and tester need to be trained on how the back end of the system operates.

#### 7.4.14 RESPONSIBILITIES

Table 13: Development and Testing Phase Responsibilities

Tasks	TM	PM	Dev team	Test team	Client
Acceptance Test Documentation & Execution	✓				✓
System/Integration Test Documentation & Execution	✓		✓	✓	
Unit Test Documentation & Execution	✓		✓	✓	
System Design Reviews	✓	✓	✓	✓	
Details Design Reviews	✓	✓	✓	✓	
Test Procedures And Rules	✓				
Screen & Report Prototype Reviews	✓		✓	✓	
Change Control And Regression Testing	✓	✓	✓	✓	

#### 7.4.15 SCHEDULE

This section contains the overall testing phase schedule of ROUTE-TO-PA. It provides the phases and key milestones as they relate to quality assurance. Furthermore, it discusses the testing goals and standards that we will like to achieve for each phase of testing that will be deployed, e.g. Unit test, Integration test, acceptance test and regression test

Table 14: Rough Schedule of Testing Phases

No	Task Name	Start	Finish	Effort	Tester	Comments
1	Unit Testing				Tester/Developer	
2	System Testing				Tester/Developer	
3	Acceptance Testing				Testers/Client (Users)	
4	Regression Testing				Tester	
5	Test Summary Report				Test Manager/ Tester	

#### 7.4.16 PLANNING RISKS AND CONTINGENCIES

Below are the associated possible risks identified which could possible occur during the testing phase of TET platform, also included is a contingency strategy meant for mitigation purpose of the risks. The impact of each risk identified has been rated with low, medium and high matrix in the table below.

Table 15: ROUTE-TO-PA Risk Matrix

Risk	Low	Medium	High	Contingency
TET platform quality problem that may block tests		✓		Adequate planning such as good defect triage and management will put in place while the test design will be so robust to moderate this risk.
TET platform testable features that will not install in the test environment	✓			This will be mitigated through smoke testing prior to the start of the test phase or better still as part of a nightly build or continuous integration. Also, an uninstall process will be put in place
Excessive changes made to the TET platform that invalidates test results or requires updates to test cases expected results and environments		✓		This risk will be managed through the good change process, robust test design with traceability to the requirement, user story and use case documentation.
Organisational issues such as experienced testers and communication between the test team and development team.		✓		Guideline and procedure will be provided to assist the team to walk through the phase while training will be provided for testers on how to carry out tests effectively. Also, communication tools such as skype, scrum meetings, emails, etc. will be used to promote effective communication.

#### 7.4.17 APPROVALS

The test plan will be approved by the following personnel: Testers, Tester Manager and Unit Manager

Date	Name	Position / Role	Remarks
21/01/16	Samuel Adebayo	Tester	
21/01/16	Ed Osagie	Test Manager	
21/01/16	Adegboyega Ojo	Unit Manager	

#### 7.5 SYSTEM VALIDATION

Similar to the manner of the definition of the Verification in software engineering, the commonest definition of Validation is the answer to the question: **“Are we building the right system?”** (Easterbrook, 2010). Validation is another set of the software product development jobs described which, in a layman’s language, refers to the process of giving official sanction, confirmation or approval to something usually through checking in order to prove the accuracy or otherwise of one thing against present attributes. In agreement with Baresi & Pezze (2006). Easterbrook’s diagram (Easterbrook, 2010) also demonstrates the fact that verification activities are not part of the activities that are involved in creating standardised requirement specifications instead the validation activities.

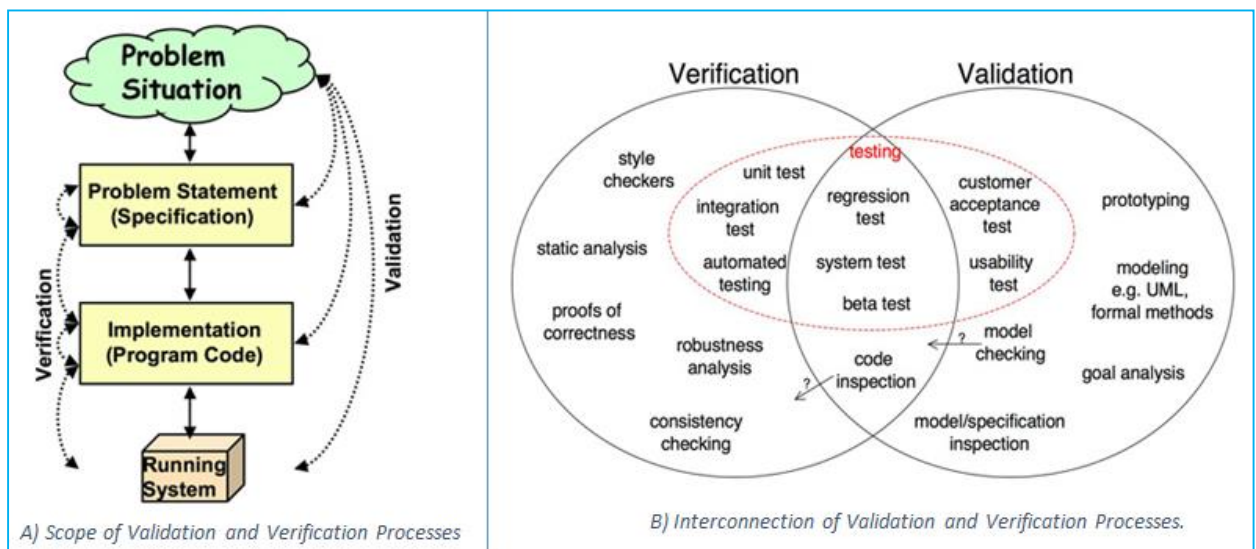


Figure 22: A) V&V Scope, B) V&V Interconnections. (Easterbrook, S. (2010))

Validation process involves making assessments (through analysis of requirements) of how well the proposed system addresses a real-world need. It includes activities such as requirements modelling, prototyping and user evaluation based on their needs and constraints. The following are also the Validation processes in V-Model: Based on Baresi & Pezze (2006), **Unit tests designed** in the module design phase can be executed either in the module design phase or executed during the validation phase whereas **Integration Testing**, which is associated with the architectural design phase, are performed to test the coexistence and communication of the internal modules within the system at the Validation phase. From the (B), Easterbrook, 2010 places both testing exercises under verification process while **User (Customer) Acceptance Test** and **Usability Test**



are performed under the Validation process. What is common to both Verification and Validation processes are code inspection and testing high-level activities such as **Regression, system** and **Beta Testing**.

While **System Testing** is directly associated with the System design phase, the aim is to check the entire system functionality for error-free quality and the communication of the system under development with external systems. It is expected that most of the software and hardware compatibility issues can be discovered during system test execution. On the hand, **Acceptance Testing** examines the business requirement analysis phase and tests the product in the user environment; the objectives of the test include discovery the compatibility issues with the other systems available in the user environment and the non-functional problems such as load and performance defects in the actual user environment (Baresi & Pezze, 2006).

## 7.6 DISTINCTION BETWEEN VERIFICATION AND VALIDATION

In Easterbrook's<sup>13</sup> blog, he distinguished between Verification and Validation processes in software development domain. These are the main aspects examined by the author:

- “Validation is concerned with checking that the system will meet the customer’s actual needs, while Verification is concerned with whether the system is well-engineered, error-free, and so on”.
- Verification will help to determine whether the software is of high quality, but it will not ensure that the system is useful to the users because varication does not look at the user needs analysis before the point when requirement specification is stablished.

This means that the distinction between the Validation and Verification is hinged on the impact of requirement specification because Validation process seeks to find whether the specification captures the customer’s needs while verification process seeks to prove that the software produced meets the specification. Consequently, if specifications are wrongly captured through the validation process, the verification process can still be right but the software produced will not meet user needs.

An important idea gained from Easterbrook’s blog for application in the ROUTE-TO-PA project is the update to connotation in the traditional software lifecycle, when “...verification [was] often taken to mean checking that the products of each phase satisfy the requirements of the previous phase [and] Validation [process] is relegated to just the beginning and ending of the project: requirements analysis and acceptance testing. He argues further that the above position assumes that the users’ requirements can be captured completely at the start of a project and that those requirements will not change development lifecycle. Our point of agreement with this idea is that, in practice, the requirements do change throughout a project, partly in reaction to the project itself and unforeseen circumstances or constraints. With this in mind, the update is that both validation and verification activities are needed throughout the lifecycle of software production.

## 7.7 USER GUIDE – TET COMPONENTS

This section is concerned with providing potential end-users (acting in this case as external testers) a set of instructions to test the functionalities of TET interface as a subsystem of ROUTE-TO-PA system, the second subsystem of the project (i.e. SPOD) is being handled by another member of the consortium.

**Testing and Recording Observations:** Please read each set of instructions given to each component to be tested before starting to execute the test instructions. This is to enable you pre-understand the direction

---

<sup>13</sup> Easterbrook Steve, Professor, Department of Computer Science, University of Toronto, Canada.

of the instructions before you begin to click around the windows. Follow the instructions as given below each test heading (representing a TET component) and record your actual observations (actual results) against the expected results in the case of deviations using the **Comments** section of the survey tables. Please note that your comments should relate only to the context of the component you have tested and not anything else. This means that you should not add comments that do not relate to that particular component you have tested so as to avoid confusion with other components' test results. If you find no deviations from the expected results, please mark **OK** as your comment. The first part of the table contains a few more '**Direct**' questions to seek your opinions regarding your satisfaction with the functionalities of TET that you have tested one after the other. The questions featuring in this section will attempt to relate each functionality you test to the objectives of ROUTE-TO-PA system components as solutions to the needs of the users in accordance with the requirements that were gathered from stakeholders earlier in the project lifecycle. The **Solution Validation** process will use the questions embedded with **solution performance measurement metrics** to seek and measure your opinions regarding the capability of each solution component to meet one or more user needs. These needs are expressed in accompanying scenarios in the forms that:

- support for **Data Transparency** for users
- facilitate **Data Accessibility** to users
- enable users to (pre-) **Check the Metadata** elements of datasets
- enable users to view the **Measure and Visualisation the Completeness of the metadata** provided for each dataset on the portal by the suppliers
- enable users to **Personalise Search criteria for Datasets** (while Logged in and not logged into the system)
- enable users to **Request Recommendation for Datasets**
- enable users to **Linking Datasets together**, etc.

## 7.7.1 TET COMPONENTS TESTING INSTRUCTIONS

### 7.7.1.1 TEST 1: CHECK METADATA QUALITY AND CHECK METADATA COMPLETENESS

**Assumed Scenario:** You are a user in need of a specific dataset for planning purposes. You have no time for repeating the job of planning and cannot afford to have a data-driven plan that does not meet the needs of the users of the plan. You want to check on the metadata quality of some datasets of interest to you (one by one) to see if and all the datasets of interest meet the metadata quality specification for the use into which you intend to put them e.g. linking a couple of datasets together before analysing them to gain insight necessary to support planning report.

**Test Data:**

- Url: <http://vmdatagov01.deri.ie:8080/dataset>
- Dataset name: **Journey Times 11122015 13.01**

**Actions and Expectations/Results/Screen displays:**

Type the url on your browser and press enter to open the TET Platform.

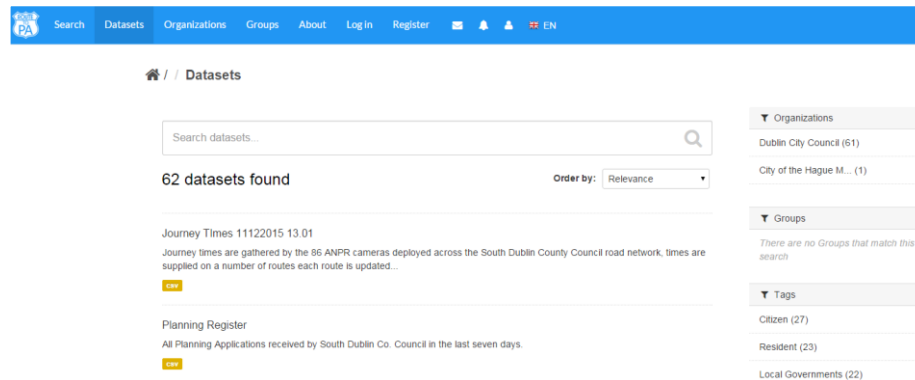


Figure 23: Check metadata quality and check metadata completeness - Screenshot 1

You should see on your screen:

A number of available datasets (**xx datasets found**) and, **Order by** (drop-down menu) field: *Relevance, Name Ascending, Name Descending* and *Last modified*.

Click on the dataset name **Journey Times 11122015 13.01** or any other dataset name on the page. The system should display the screen similar to the one below.

## Journey Times 11122015 13.01

Journey times are gathered by the 86 ANPR cameras deployed across the South Dublin County Council road network, times are supplied on a number of routes each route is updated every 5 minutes in csv format.

### Data and Resources

**Journey Times 13.01**  
Journey times are gathered by the 86 ANPR cameras deployed across the South...

[Explore](#)

Population Transport and Infra... data scientist journey-times traffic traffic-management

### Additional Info

Field	Value
Language	English
Target Audience	data scientist
Category	Transport and Infrastructure
Related Datasets	traffic-cameras
Source	<a href="http://dublinked.com/datastore/datasets/dataset-260.php">http://dublinked.com/datastore/datasets/dataset-260.php</a>
Version	1.1
Frequency of Publication	Every 5 minutes
Was Revision Of	<a href="http://dublinked.com/datastore/datasets/dataset-260.php">http://dublinked.com/datastore/datasets/dataset-260.php</a>
Spatial Coverage	<a href="http://www.sdcc.ie">http://www.sdcc.ie</a>
Spatial Notes	Boundary box ((53.166000,-6.556000),(53.381000,-6.259000)) Projection ITM
Temporal Coverage	
Temporal Coverage From	
Temporal Coverage To	
Temporal Notes	
Conforms to	<a href="http://w3.org/TR/dwbp/">http://w3.org/TR/dwbp/</a>
Author	GIS Officer SDCC
Author Email	<a href="mailto:GIS Officer SDCC">GIS Officer SDCC</a>
Maintainer	Margaret Twynam Muldoon
Maintainer Email	<a href="mailto:Margaret Twynam Muldoon">Margaret Twynam Muldoon</a>
Completeness	85 %

Figure 24: Check metadata quality and check metadata completeness - Screenshot 2

Observe the metadata fields on the left of the table and their corresponding values supplied by the data provider on the right. The metadata completeness rating is displayed at the bottom of the table as a horizontal

bar, and it measures the completeness of the metadata supplied in percentage. The more values of metadata supplied for the field, the larger the quality of the metadata and the higher is the percentage rating displayed.

Table 16: TET User Survey - Test 1

Questions	Responses
1. Does this functionality seem good enough to improve accessibility to datasets?	
2. Will the possibility for you to see the metadata of the dataset beforehand affect your decision to use or not to use the dataset?	
3. What value does this system capability add to the existing open data platform?	
4. How would you rate your satisfaction with the manner of implementation of this functionality? Choose from 1-10. (1 = no satisfaction, 5=good, 10 = excellent)	
Additional Comments	

#### 7.7.1.2 TEST 2: PROFILE-BASED PERSONALIZATION (FOR A USER WHO HAVE NOT LOGGED IN)

**Assumed Scenario:** As a user, you are not interested in creating a profile on the platform or who does not want to log in at this moment, but you are interested in searching for datasets that meet a given sets of criteria, perhaps different from what were earlier saved on your profile account. You provide the system with a defined set of criteria by completing a search query form and clicking on the search button afterwards. In this scenario, the search criteria are not saved on the system and can only be repeated through the same process as before. On the clicking the **search** button, you should retrieve from the system, the available datasets that match the criteria you provided.

**Test Data:**

**Url:** <http://vmdatagov01.deri.ie:8080/>

**Dataset:** Data Scientist or any other datasets by Role category, or by any other category.

**Actions and Expectations/Results/Screen displays:**

Go to the given url on your browser; the system opens the page below and display datasets under **Role (I am...)**; shows the remaining categories e.g. **Datasets by Category** and **Datasets Recommended for me**

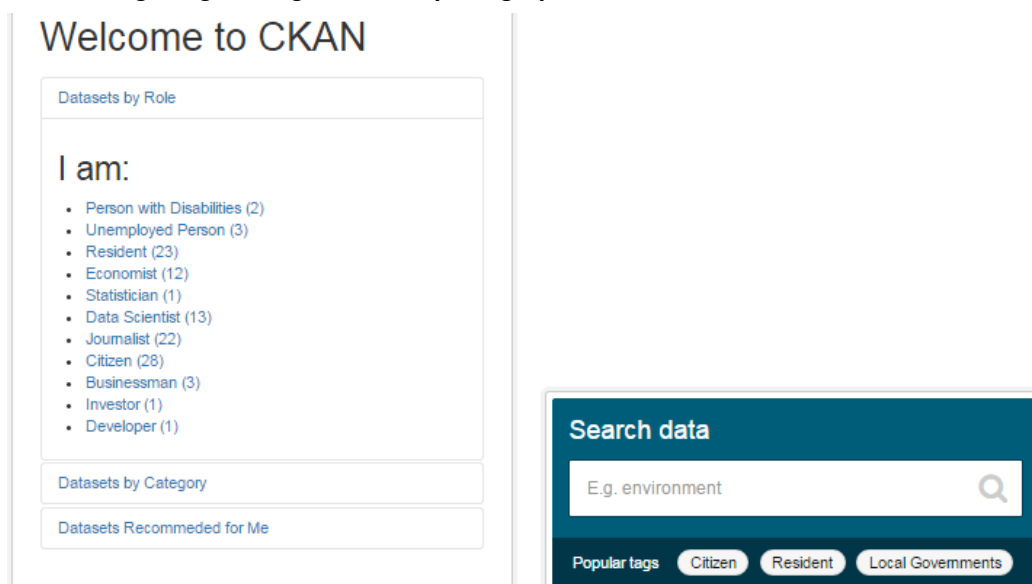


Figure 25: Profile-based Personalization (user not logged in) - Screenshot 1

Click on the dataset name specified above (or another of your choice from the list) and system should display the screen below or something similar:

# Datasets for Data Scientist

## Fingal Open Data Airport

Data relating to the airport in Fingal County Ireland

csv xml xml

## Freedom of Information Request

List of number of requests for information and where they came from, dates etc. This dataset does not contain the nature of the request.

Txt xlsx

## Journey Times 11122015 13.01

Journey times are gathered by the 86 ANPR cameras deployed across the South Dublin County Council road network, times are supplied on a number of routes each route is updated...

csv

Figure 26: Profile-based Personalization (user not logged in) - Screenshot 2

Alternatively, you may personalise your search by using the search box provided on the first screen when the url is entered into the browser.

Type in any search word/tag e.g. **“Economist”** and press **enter key**. The system will return the number of datasets matching the search tag and displays the filenames. If the number found is more than can be displayed on one page, the system will display them on multiple pages. Scroll down the screen to view the multiple pages and select the pages one by one to view the datasets filenames.

Table 17: TET User Survey - Test 1

Questions	Responses
1. Does this functionality appear good enough to improve user accessibility to the relevant datasets?	
2 How relevant and reliable are the outcomes produced by using this feature?	
3. How would you rate your satisfaction with the manner of implementation of this functionality? Choose from 1-10. (1 = no satisfaction, 5=good, 10 = excellent)	
Additional Comments	

**Continuation of Non-logged in Personalized Search** – This is an alternative way to achieve **Personalized Search** without logging into the TET platform.

**Test Data:** <http://srvgal100.deri.ie:8081/>

**Dataset:** Person with Disability

**Actions and Expectations/Results/Screen displays:** Enter the url on you browser and press **Enter**. The TET home page opens showing the Welcome Message. Scroll to the bottom of the page and observe various icons under the heading **“What are you looking for?”** Each icon represents a category of datasets.

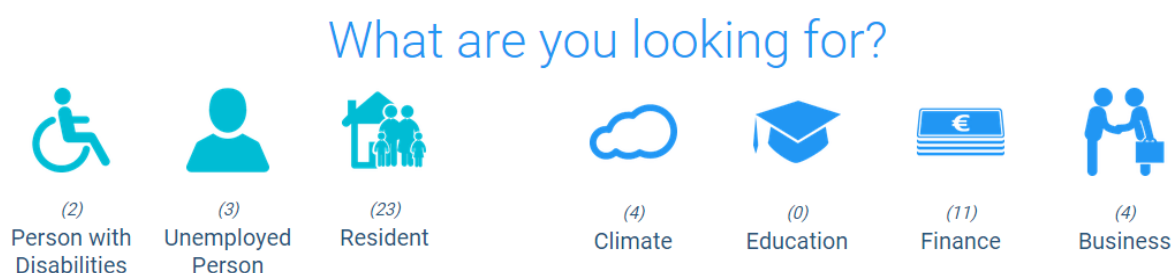


Figure 27: Profile-based Personalization (user not logged in) - Screenshot 3

Click on the dataset icon stated above or select any one of them which is of interest to you. The system should display the number of datasets found in that category you selected and the datasets file names.

Table 18: TET User Survey - Test 1

Questions	Responses
1. Does this functionality appear good enough and improves the simplicity of searching for relevant datasets?	
2 How relevant is this section of the systems to the speed of searching the platform?	
3. How would you rate your satisfaction with the manner of implementation of this functionality? Choose from 1-10. (1 = no satisfaction, 5=good, 10 = excellent)	
Additional Comments	

#### 7.7.1.3 TEST 3: PERSONALIZED SEARCH (USER IS LOGGED IN WITH ADMIN RESPONSIBILITY)

**Assumed Scenario:** As a regular open data provider or consumer, you are perhaps public sector personnel, a professional Data Journalist, a Data Scientist, a community Leader/Activist, who is always interested in government open data; an Open Data System Developer or any other data consumer. You have registered your profile and data interest on the platform so that datasets retrieval for your frequent uses is done seamlessly and more accurately.

**Url:** <http://vmdatagov01.derri.ie:8080/user/login>

**Dataset:** User details with an administrative capacity – Use the following:

- User name: **admin**
- Password: **\*\*\*\*\***

**Actions and Expectations/Results/Screen displays:** Go to the url above and enter admin login details provided on the user account form that opens, then click on **Login** button. Note, you may tick on the *Remember me* box for auto login next time you visit the systems.

Figure 28: Personalized search (user login: admin responsibility) - screenshot 1

Once the Login button is clicked, the Dashboard screen opens showing: **News field** (opened by default), **My Datasets**, **My Organisation**, and **My group**. The **edit** button is shown on the right side of the screen.

Figure 29: Personalized search (user login: admin responsibility) - screenshot 2

Click on “**My Datasets**” to display the available datasets under the current admin login. There is an “**Add Dataset**” button on top of the list of datasets displayed on the screen which the user should use to add more dataset to the system. The user can click on any dataset file name to view the metadata and the completeness bar of the metadata of a dataset.

Table 19: TET User Survey - Test 1

Questions		Responses
1. Do you consider this functionality to be well personalised and user-friendly?		
2 How accurate do you think the result of the personalised is?		
3. How would you rate your satisfaction with the manner of implementation of this functionality? Choose from 1-10. (1 = no satisfaction, 5=good, 10 = excellent)		
Additional Comments		

#### 7.7.1.4 TEST 4: PERSONALIZED SEARCH (USER IS LOGGED IN)


**Assumed Scenario:** Assume again in this scenario, you’re a regular open data consumer (but not a provider), who is perhaps in the profession of Data Journalism or as a Data Scientist, a Community Leader/Activist who is always interested in government open data; an Open Data System Developer or any other data consumer. You have registered your profile and data interest on the ROUTE-TO-PA platform so that datasets retrieval for your frequent uses is done seamlessly and more accurately.

**Url:** <http://vmdatagov01.deri.ie:8080/user/register>

**Dataset:** For User details as a consumer of datasets capacity (a common portal user), use the following:

- User name: **tetuser**
- Password: **\*\*\*\*\***

**Actions and Expectations/Results/Screen displays:** Go to the url. above and click on **Register** button. The system should display the form below:

 / / Registration

## Register for an Account

Username:

Full Name:

Role:

Category of Interest:

Twitter ID:

Email:

Password:

Confirm:

[Create Account](#)

Figure 30: Personalized search (user is logged in) - screenshot 1

In the form that opens type in the details above and complete the rest of the fields as you wish. Note that compulsory field must be completed else the system will not save your details. When done, click on **Create Account** button.

In the screen, that displays, click on **Datasets** button on the top menu.

The system should display a list of datasets available on the platform repository and **Personalized Search** button for the user to filter the result based on his/her user profile data.

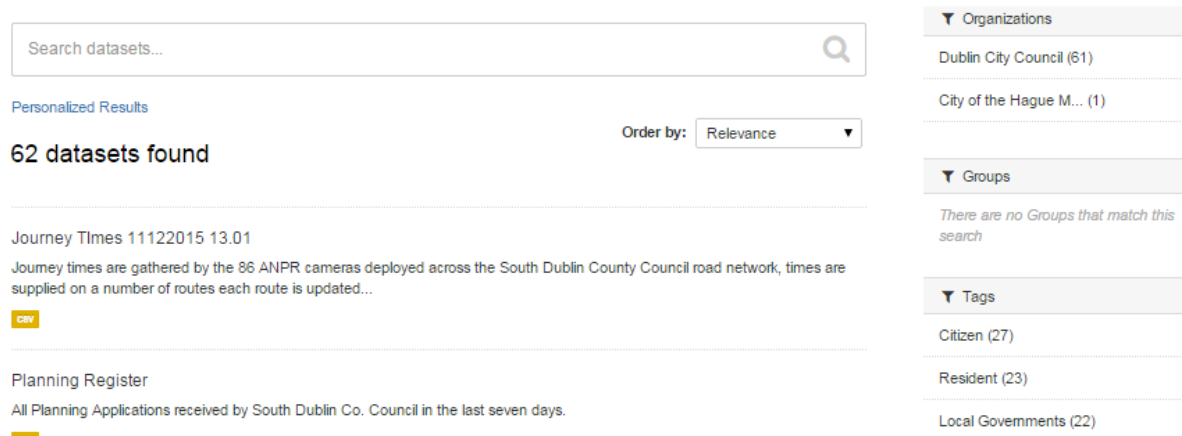


Figure 31: Personalized search (user is logged in) - screenshot 2

Click on the **Personalized Search** button to display the result of datasets matching the user's profile and interests.

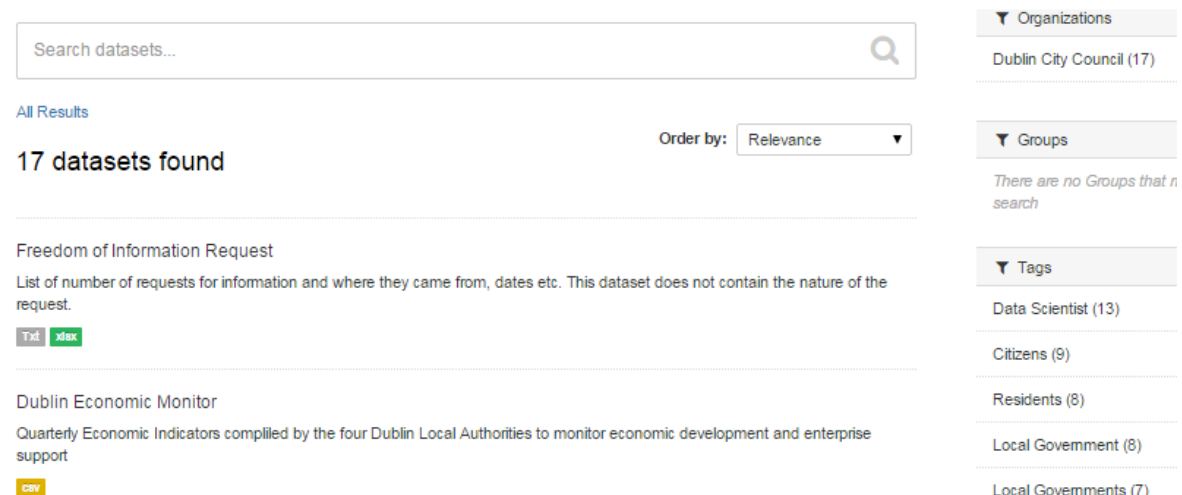


Figure 32: Personalized search (user is logged in) - screenshot 3

Observe that the **Personalized Search** button has changed to **All Results** button and the number of datasets displayed is reduced according to the matches found. By clicking on the **All Results** button, you can toggle between the total datasets displayed and the number of datasets matching the user profile. On the right side of the screen are tags and headings to filter the search results.

Table 20: TET User Survey - Test 1

Questions	Responses
1. Is the searching based on personalised profile account user –friendly, seamless and fast?	
2 How accurate are the results of the searches?	
3. How would you rate your satisfaction with the manner of implementation of	



this functionality? Choose from 1-10. (1 = no satisfaction, 5=good, 10 = excellent)	
Additional Comments	

#### 7.7.1.5 TEST 5: PERSONALIZED RECOMMENDATIONS

**Assumed Scenario:** This scenario is similar to the one in which the user needs to use personalised search functionality so that the system can generate the matching datasets based on the user preferred profile data. In addition to this, in the case of **Personalised Data Recommendations**, the user enters more details such as Datasets category of interest, email address, the Role of the user e.g. Data Scientist, Unemployed person, Economist, Disabled person, etc. In the future, there will be a possibility to add a choice of email alert for users to receive an update to datasets or a fresh upload in their areas of interest.

**Url:** <http://vmdatagov01.deri.ie:8080/>

**Dataset:** Planning Register.

**Actions and Expectations/Results/Screen displays:** Go to the url specified above; ensure that you are not logged into the system yet. Then locate the button “**Datasets Recommended for me**” at the bottom of the box displayed

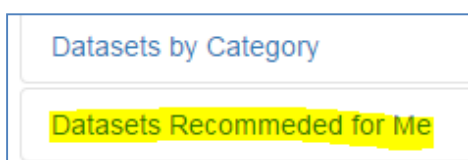


Figure 33: Personalized recommendations - screenshot 1

Click on the button and observe the datasets displayed by the system. You may take a screen shot of the screen to compare with the next result below. The datasets displayed on this occasion (not logged in) are based on the popular datasets searches by users.

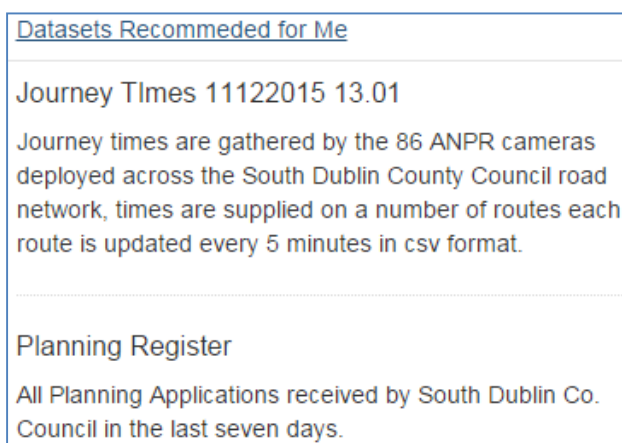


Figure 34: Personalized recommendations - screenshot 2

Now, log into the system: Click on the **Login** button on the top menu and enter the required user (your user details). Once logged in, click or the **Search** button on the left of the top menu. That should bring you back to the screen initial screen where the button “**Datasets Recommended for me**” is located. Click on the “**Datasets Recommended for Me**” again (this time, while you are logged into TET systems):

<a href="#">Datasets Recommended for Me</a>
<p><b>Freedom of Information Request</b></p> <p>List of number of requests for information and where they came from, dates etc. This dataset does not contain the nature of the request.</p>
<p><b>Dublin Economic Monitor</b></p> <p>Quarterly Economic Indicators compiled by the four Dublin Local Authorities to monitor economic development and enterprise support</p>

Figure 35: : Personalized recommendations - screenshot 3

Observe that the datasets displayed in the while you are logged into the system are slightly different from those displayed while you were not logged into the system. The datasets recommended while logged in are personalised based on your profile data while those recommended while not logged in are not based on your profile details instead on the popular searches of general users.

**Note:** In the future version, the system will be configured so that users who want dataset recommendation would be able to choose if they would want data update or upload alert sent to them via emails so that they have quick notice of the availability of new datasets or updated versions for their immediate uses.

Table 21: TET User Survey - Test 1

<u>Questions</u>	<u>Responses</u>
1. Do you think this is a relevant functionality?	
2 How well do you think this feature will facilitate open data use and platform adoption?	
3. Would you use this tool to request data recommendations?	
4. How would you rate your satisfaction with the manner of implementation of this functionality? Choose from 1-10. (1 = no satisfaction, 5=good, 10 = excellent)	
Additional Comments	

#### 7.7.1.6 TEST 6: RECOMMEND RELATED DATASETS

**Assumed Scenario:** This functionality follows on from the other ones already treated above. It uses the datasets types already matching your criteria at any occasion (logged in or not logged in) to recommend further dataset that would possibly be of interest to you. In the case of logged in, the related datasets would be those datasets that are similar in category or nature to the ones much more closely matching your profile data whereas when not logged in, the related datasets would be those datasets coming next after the most popular datasets being consumed by the general user community.

**Url:** <http://vmdatagov01.deri.ie:8080/dataset>

**Dataset:** Planning Register, Dublin City Libraries Accessibility Audit or Sculpture in Dublin City Council Parks

**Actions and Expectations/Results/Screen displays:** Open the url above:

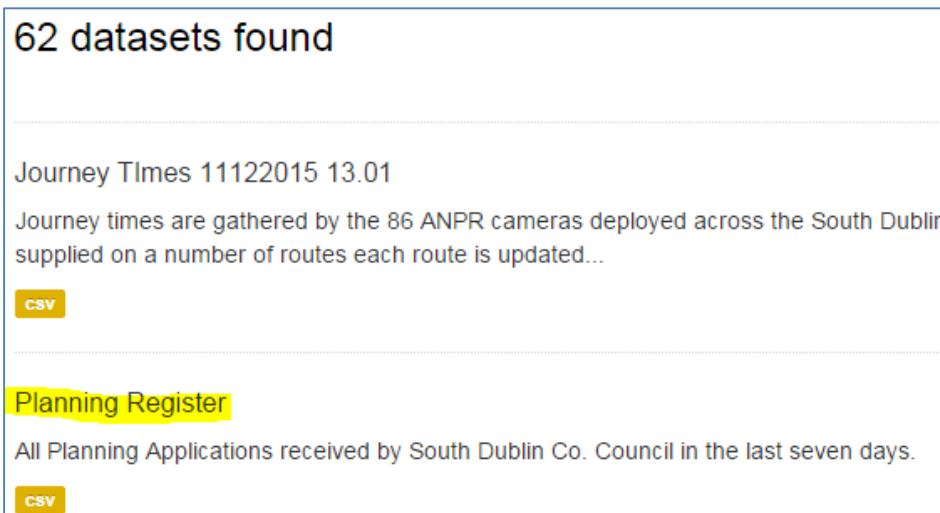


Figure 36: Recommend related datasets - screenshot 1

Then click on the named dataset or any other dataset you can find in the page to view the metadata of the dataset. Scroll to the bottom of the metadata table, then locate the dataset(s) displayed under the heading “**Related Datasets**”. Click on the dataset name to view the related dataset(s) displayed by the system.

## Planning Register

All Planning Applications received by South Dublin Co. Council in the last seven days.

### Data and Resources

#### Planning Base

All Planning Applications received by South Dublin Co. Council in the last...

Explore

Preview

Download

Land Use and Zoning Public applications land-use planning

Figure 37: Recommend related datasets - screenshot 2

You can explore the dataset by using the drop down button **Explore** to give you access to actions - preview or download of the dataset. Other search tags are available to trace the related datasets e.g. Public, Land-use, etc.

**Note:** There are some datasets that do not at this version have related datasets to display when you click on the filename. Example: Customer service request (fixyourstreet.ie) has no related dataset yet on this version of TET portal.

Table 22: TET User Survey - Test 1

Questions		Responses
1. Do you think the automatic display of related datasets to the ones you have searched for will improve dataset visibility and accessibility?		
2 How relevant is this feature considering the already available recommendation of the dataset and the personalised search functions?		
3. How would you rate your satisfaction with the manner of implementation of this functionality? Choose from 1-10. (1 = no satisfaction, 5=good, 10 = excellent)		
Additional Comments		

#### 7.7.1.7 TEST 7: DATASET LINKING

**Assumed Scenario:** Note that in this version; only portal administrators can use this functionality because it is not yet available for ordinary data consumers. Therefore, backend administrator's login detail is required as shown below:

- User name: **admin**
- Password: **\*\*\*\*\***

A subsequent version will be incorporated into front-end version of the tools for linking dataset and made available to data consumers.

As a data provider, assume that you have the need to inform or enable data users that one or more datasets are related to another by whatever nature. Perhaps the latter dataset is an update to the former and would make a better meaning by using both together or understanding something about the latter version would improve the use into which the former or both datasets can be put. Linking datasets will enable the system to recommend the other dataset(s) as related datasets when one of them is searched out by a user because linking functionality treats datasets as related. Secondly, linking dataset enables both datasets to be integrated into one dataset (*future version of TET*)

**Url:** <http://srvgal100.deri.ie:8081/>

**Dataset:** Dublin City Libraries Accessibility Audit and Planning-register datasets

**Actions and Expectations/Results/Screen displays:** Open the given url and arrive at the home TET page. Then click on the login icon on the top right corner of the page. The system should display the login form.

Complete the login field with the admin data provided above and you may tick-off the box for **Remember me** option before clicking on the **login** button.

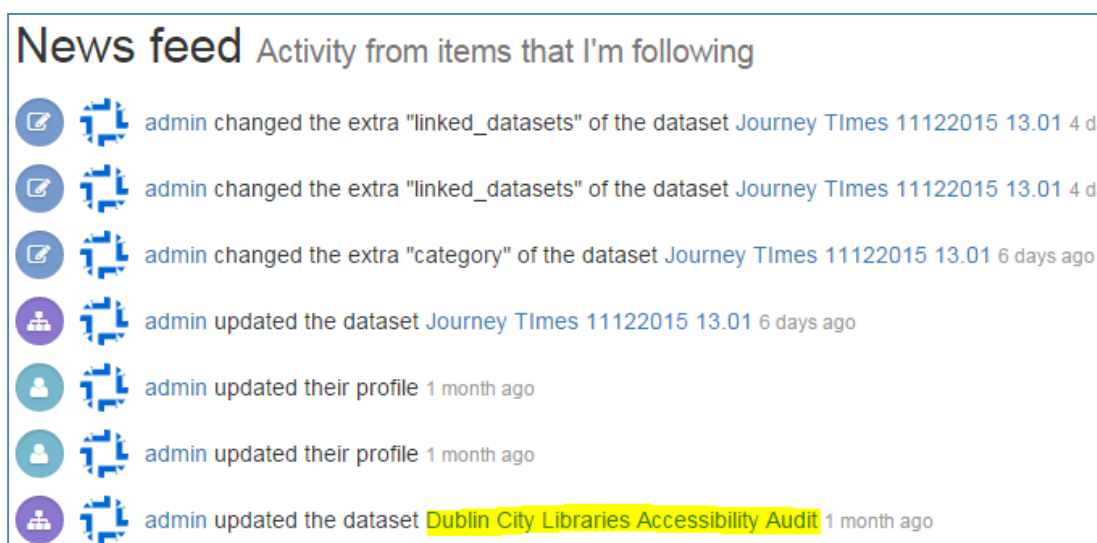



Figure 38: Link dataset - 1

Click on the given dataset name as seen on the highlighted portion of the screenshot above (Not you can select any dataset of your choice). On the screen that opens, locate the **Manage** dataset button at the middle of the screen, and then click on it and the system should open the **Edit Metadata** table.

\* **Category:** × Finance

**Related Datasets:** planning

\* **License:** planning-register

 License definitions and additional information can be found at <http://opendefinition.org/>

**Organization:**

Figure 39: Link dataset - 2

In the table, locate the Related Datasets field, and as you begin to type in the name of the dataset (Planning-register) to be linked with the first (Dublin City Libraries Accessibility Audit), you get name suggestions. Click on the name required to complete the action. When done, click on the **Update Dataset** button at bottom right of the page.

The system should return to the metadata table for the first dataset (i.e. Dublin City Libraries Accessibility Audit). On this screen observe that for the field **Related Datasets**, the value is the linked dataset.

<b>Category</b>	Finance
<b>Related Datasets</b>	planning-register

Figure 40: Link dataset - 3

#### To unlink a dataset from another:

From the metadata screen of a linked dataset, click on the **Manage** button arrive at the **Edit Metadata** table. Go to the **Related Datasets** field and just click on the **x** button on the linked dataset to remove it. Click on the **Update Dataset** button to save.

Table 23: TET User Survey - Test 1

<b>Questions</b>		<b>Responses</b>
1. What importance does this functionality serve to you?		
2. Do you find the performance of this functionality efficient and usable?		
3 Does this functionality enhance the traceability of datasets and hence availability for users' consumption?		
3. How would you rate your satisfaction with the manner of implementation of this functionality? Choose from 1-10. (1 = no satisfaction, 5=good, 10 = excellent)		
Additional Comments		

## 8 CONCLUSION

---

The release of ROUTE-TO-PA Alpha version of TET was built based on the information that was documented in the previous deliverables such as D2.2, D2.3 and D2.4 earlier highlighted in this project. The goal of this alpha version release of TET is to achieve one of the important objectives of ROUTE-TO-PA, which is to provide tools that could be integrated into existing open data platforms to deliver greater data transparency, quality and understandability. In this deliverable - D4.2 Alpha version of TET - we have used the system requirement gathered from the requirement analysis as a guide in the development phase of the platform. Furthermore, the system requirements served as the basis for developing the architectural design of TET model, which consist of TET architecture and the implementation technology. Moreover, this deliverable documents the deployment requirement and process, which identify the preferred environment for TET: setting up the server and installation of the relevant TET components. Additionally, it also highlights the configuration process of the TET platform which enhances the capability of the platform to be either fully integrated as an open data solution or some components of it can be adapted to suite existing portals.

Finally, in the verification and validation section, we have provided detailed information about how the testing phase of the platform will be carried out. The information provided includes a test framework to explain the chosen approach adopted supported by a test plan, which narrates the procedures that serve as a guide to the test activities. Included in this section is also a user guide with instructions to guide potential end-users of the platform during the evaluation test of the implemented functionalities of TET components as part of TET subsystem of ROUTE-TO-PA.

In a subsequent version of TET subsystem, more components will be implemented and also supplied with the user guides to enable evaluation processes to take place. Before then, the implementation team hopes to receive evaluated report of the Alpha version so that early corrections of issues and bugs found by solution evaluators can be carried out. It is hoped that the next version – Beta version will be released by January 2017.

## 9 APPENDICES

---

### TEST SPECIFICATION (TEST CASES)

The ROUTE-TO-PA test cases described hereunder were created from the corresponding use case that was in turn based on the user stories and the system requirements. The tagging (ID) of the test case directly suggests the relevant use case it relates e.g. Test Case 1 (TC1) based on Use Case 1 is tagged TC1-UC1. Where multiple test cases were produced from one use case, we introduced decimal numbering into the tagging system. For example, test case ID, TC2.1-UC2 refers to the first test case produce in use case 2 and TC.2-UC2 refer to the second test case based on use case 2 also.

The following tables are the test cases based on the use cases in deliverable D2.4 (Requirement Specification and Use Case Models). It is worth mentioning that in the test case table, the use case flows were repeated just for the refreshment of the mind and most of the test data for testing the functionalities are unavailable at this stage. However, the required test data will be produced and added to subsequent revisions of the test case documentation until the final release of the beta version of the product.

Table 24: Check Metadata Quality and Completeness

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC1-UC1	Check Metadata Quality and Completeness	Dataset supplied to the portal must meet given quality properties. This test will verify in the graphical manner the completeness of all metadata fields for that dataset	CKAN Platform CKAN dataset Catalogue url A dataset to be checked for completeness	1. Horizontal bar showing completeness of metadata qualities 2. the various metadata qualities	A dataset available on CKAN dataset catalog
Step #	Use case steps	Test Procedure / Steps	Expected Results (I = Internal)	Actual Result	Pass/Fail
1	User selects a dataset in the dataset catalogue of CKAN	Click on "Dublin Traffic 2014" Dataset	System selects the dataset and reads the metadata details (I)	NR	NA
2	None	None	System displays 1. Metadata qualities - the name of the supplier, the name of contact person, email of contact person, date last updated, provenance record, context description. 2. Visualised completeness bar calibrated in percentages: 10, 20, 30...100%	NR	NA
3	User selects one metadata quality	Check on one metadata quality Checkbox	Quality status (Completeness) bar should increase in length if there is information on the quality field	NR	NA
4	User selects multiple metadata qualities	Check on multiple metadata quality Checkboxes	Quality status (Completeness) bar should increase in length if there is information on the quality fields	NR	NA
5	User deselects multiple metadata qualities	Check off multiple metadata quality Checkboxes	Quality status (Completeness) bar should decrease in length if there is information on the quality fields	NR	NA



Table 25: Add Provenance to Dataset

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC2.1-UC2	Add Provenance to Dataset	This test is to demonstrate how system should respond to adding provenance record to dataset before download	To test this functionality, Tester needs to 1. Open the CKAN platform 2. Open the CKAN data catalogues 3. Select a dataset which is to be downloaded or used in other ways on platform.	1. System must include provenance records with the dataset downloaded 2. The system must present the content along with the dataset to the tester	Dublin Traffic 2014 Dataset
Step #	Use case steps	Test Procedure / Steps	Expected Results (I = Internal)	Actual Result	Pass/Fail
1	User clicks on the dataset of interest to the user	Click on "Dublin Traffic 2014" Dataset	System displays the dataset	NR	NA
2	User checks on the ADD PROVENANCE button	click ADD PROVENANCE button	System reads provenance field and displays details on screen	NR	NA
3	User clicks on SAVE button to save dataset (optional)	Click on SAVE button (optional)	System saves the dataset	NR	NA

Table 26: Add Provenance to Dataset (Exception - Not Provenance records on dataset)

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
T2.2-UC2	Add Provenance to Dataset (Exception - Not Provenance records on dataset)	This test is to demonstrate how system should respond to having no provenance record on a dataset	To test this functionality, Tester needs to 1. Open the CKAN platform 2. Open the CKAN data catalogues 3. Select a dataset	System display a warning message about lack of provenance record	Dublin Traffic 2014 Dataset
Step #	Use case steps	Test Procedure / Steps	Expected Results (I = Internal)	Actual Result	Pass/Fail

1	At step 2 of norm of T2.1-UC2: User clicks on ADD PROVENANCE button	At step 2 of norm of TC2.1-UC2: clicks on ADD PROVENANCE button	System displays warning message [Provenance Records not available]	NR	NA
2	3. User clicks on OK button	Click OK button	System returns to step 2 of the normal flow	NR	NA

Table 27: Check and rate dataset quality [View Dataset quality]

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC3.1-UC3	Check and rate dataset quality [View Dataset quality]	Test for the check and rate functionality of the dataset quality on the platform	1. User must have login details, 2. User must be logged on to the platform and be in the data catalogue window	System should display the various quality parameters of the dataset including metadata, provenance, context, date of last update, details of the publishers, and number of missing field.	Dublin city Traffic 2014 Dataset
Step #	Use case steps	Test Procedure	Expected Result	Actual Result	Pass/Fail
1	User select the dataset of interest	Click on the dataset of interest [Dublin City Traffic 2014]	System displays dataset and VIEW QUALITY button	NR	NA
2	User VIEW QUALITY button	click the VIEW QUALITY button	System displays the dataset quality parameters and a BACK navigation button	NR	NA
3	User views the dataset qualities, then clicks on BACK button	click on the back button to return	System returns to the previous page at step 2	NR	NA

Table 28: Check and rate dataset quality [Rate dataset quality]

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC3.2-UC3	Check and rate dataset quality [Rate dataset quality]	Test demonstrates the functionality for rating or ranking dataset quality by user	1. User must have login details, 2. User must be logged on to the platform and be in the data catalogue window	System should enable the user to rate the qualities of the dataset he/she views on the platform in a simple manner	Dublin city Traffic 2014 Dataset
Step #	Use case steps	Test Procedure	Expected Result	Actual Result	Pass/Fail
1	At step 2 of T3.1, user clicks on RATE button	Click on the RATE button	System presents a scale to rank the dataset and a small textbox to describe the dataset quality [word length to be determined]	NR	NA
2	User rates the dataset quality	1. Rank the dataset quality 2. write on the textbox (optional) about the dataset quality	1. System saves the rating, 2. System displays a notification message to show completeness 3. OK button	NR	NA
3	User clicks OK button	click the OK button	System returns to the Dataset quality screen at step 2	NR	NA
4	None	None	system present a dialogue box with message: 'create another entry'; an OK and DONE buttons	NR	NA
5	user clicks on OK button	click the OK button	system present a pop-up window with dropdown menu of widgets	NR	NA
6	User repeats steps 2 - 4 as desired	Repeat steps 2 - 4 as desired	System repeats steps 2 – 4	NR	NA
7	user clicks on DONE button	click on the DONE button	system display all entries and a dialogue box options: ADD, EDIT, POST	NR	NA
8	user clicks on POST	Click on POST/SEND button	system send post to network participants and display notification		

Table 29: Link a Dataset

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC14-UC14	Link a Dataset	This test will verify the system capability to link datasets together whether from the external repository or within the ROUTE-TO-PA facility	User login and a dataset a necessary to carry out this test	System should link datasets successfully together according to user need and format	1. User login details 2. "Food Hygiene Information Scheme Rating Glasgow" dataset
Step #	Use case steps	Test Procedure / Steps	Expected Results (I = Internal)	Actual Result	Pass/Fail
1	User selects a dataset to be linked	Click on "Food Hygiene Information Scheme Rating Glasgow" dataset	System displays: a. the dataset and b. LINK and CANCEL buttons	NR	NA
2	User clicks on LINK button	click on LINK button	System displays: a. Criteria to link datasets, and b. DONE button	NR	NA
3	User defines the entities/datasets to be linked and the properties to link them	Type in the criteria and properties to link the datasets together	None	NR	NA
4	User clicks on DONE button	Click on DONE button	1. System reconciles data 2. System establishes links 3. System stores the links 4. System updates dataset with the linked data (I)	NR	NA
5	None	None	System displays: a. notification [Dataset successfully linked] b. OK button	NR	NA
6	User clicks on OK button	click on OK button	System returns to initial screen	NR	NA

Table 30 Personalize Search (User logged into account)

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC17.1-UC17	Personalize Search (User logged into account)	This test is to show how system executes personalised search for a user who has profile details on CKAN platform and is logged into account	Personal or assumed profile data of the tester must be on the system as register user on CKAN platform - while searching with login	System should apply user personal data e.g. age, occupation or role, sex, interest to search dataset for the user	User personal data: Name: Test User, Username: testuser Sex: M, Age: 50, Occupation: Developer, Interest: Photography
Step #	Use case steps	Test Procedure / Steps	Expected Results (I = Internal)	Actual Result	Pass/Fail
1	User logs into CKAN platform	Login using Username: testuser, and password: pass2015word	System opens into the testuser's account on CKAN platform	NR	NA
2	User clicks on "Personalized Search" button	Click on PERSONALISED SEARCH button	System displays search results based on the personal data of user	NR	NA

Table 31: Personalize Search (user not logged in)

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC17.2-UC17	Personalize Search (user not logged in)	This test is to show how system executes personalised search for a user who has no login or decides not to login for the search	To test this condition (i.e. personalise search with no login) user situation or search options must be available for user to choose from e.g. location, role, interest options, unemployed, jobseeker, self-employed, etc.	System should apply user-chosen option(s) to search dataset for the user e.g. interest, role, location	Sear options: Location, roles, interest (photography, weather, global warming, seaweeds, oceanography, sports) Job vacancies in Dublin area
Step #	Use case steps	Test Procedure / Steps	Expected Results (I = Internal)	Actual Result	Pass/Fail
1	User opens CKAN platform url	Open CKAN url [www.ckan.org] on browser	System opens into CKAN platform	NR	NA

2	User clicks on SEARCH button	Click on SEARCH button	System displays: 1. Notification: [Select search criteria from list displayed] 2. Search criteria: Location, Roles (on dropdown menu e.g. researcher, developer, data journalist, jobseeker, health status, etc.), level of dataset needed in a dropdown menu e.g. basic information, details, graphical, images, etc.)	NR	NA
3	User chooses required search options	Select location: Dublin; Health status: Person with disability; Level: Basic	None	NR	NA
4	User clicks on SEARCH button	Click on SEARCH button	System displays datasets (various formats) in Dublin area containing disability details	NR	NA

Table 32: Personalize Search (Exception)

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC17.3-UC17	Personalize Search (Exception)	This test is to show how exceptional case of a personalised search whereby no result is found in the first search in step 2 of T17.1-UC17	To test this condition (i.e. personalise search with no login) user situation options must be available for user to choose from e.g. location, role, interest options	System should apply user-chosen option(s) to search dataset for the user e.g. interest, role, location	Sear options: Location, roles, interest (photography, weather, global warming, seaweeds, oceanography, sports, health status) Job vacancies in Dublin area
Step #	Use case steps	Test Procedure / Steps	Expected Results (I = Internal)	Actual Result	Pass/Fail
1	User chooses "Personalized Search"	Click on PERSONALISED search button	System displays no result. [No match for your search criteria]	NR	NA
2	None	None	System displays notification [Select search criteria from list displayed]	NR	NA
3	User chooses required search option	Select location: Dublin; Role: Jobseeker	System displays job vacancies in Dublin area to match options chosen	NR	NA

Table 33: Request Recommendations for Datasets (user logged in)

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC19.1-UC19	Request Recommendations for Datasets (user logged in)	This test shows how the system can be requested to recommend dataset based on user's profile and how system presents recommendations	1. CKAN platform url 2. User login details 3. Tester should be in the dataset catalog window	System presents user with a list of recommended datasets based on user profile	1. User profile 2. List of some datasets matching profile data 3. Tester's email address: <a href="mailto:tester.user@r2pa.eu">tester.user@r2pa.eu</a>
Step #	Use case steps	Test Procedure / Steps	Expected Results (I = Internal)	Actual Result	Pass/Fail
1	User Logins into account on CKAN: enter details and clicks on LOGIN button	Enter login details and click on LOGIN button	System opens into tester's account on CKAN	NR	NA
2	User clicks on "Request Recommendations" button	Click on "Request Recommendations" button	System displays: 1. Recommendation criteria: a) Role: (dropdown menu) e.g. Jobseeker, Researcher, Developer, Teacher, etc.; b) Location; c) Email address field; d) Level of information required (dropdown menu) 2. OK button	NR	NA
3	User selects required details, and types in email address	Select Role: Developer, Location: Dublin; type in email address: <a href="mailto:tester.user@r2pa.eu">tester.user@r2pa.eu</a>	None		
4	User clicks on OK button	Click OK button	System displays: [Recommendation setting complete]	NR	NA
5	None	None	System adds user to dataset recommendations list (I)	NR	NA
6	Upload a new/updated dataset to CKAN platform (Ref to Dataset upload UC)	Upload "Dublin Traffic 2013" dataset to CKAN platform	System receives a new/updated dataset (I)	NR	NA

7	None	None	System evaluates the dataset (quality, metadata, completeness and compares it with users profiles on recommendation list) (I)	NR	NA
8	None	None	System creates a list of interested users and sends recommendation to them (I)	NR	NA
9	None	None	Tester receives dataset recommendations notification based on tester's criteria in his inbox	NR	NA

Table 34: Request Recommendations for Datasets (user not logged in)

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC19.2-UC19	Request Recommendations for Datasets (user not logged in)	This test shows how the system can be requested to recommend dataset based on user's defined criteria	1. CKAN platform url 2. Tester should be in the dataset catalog window	System presents user with a list of recommended datasets based on user-selected details	1. List of some datasets matching profile data 2. CKAN platform url 3. Tester's email address: <a href="mailto:tester.user@r2pa.eu">tester.user@r2pa.eu</a>
Step #	Use case steps	Test Procedure / Steps	Expected Results (I = Internal)	Actual Result	Pass/Fail
1	User opens into CKAN platform	Open CKAN platform url in a browser	System opens CKAN platform showing home screen	NR	NA
2	User clicks on "Dataset Catalog" button	Click on DATASET CATALOG button	System displays: 1. Dataset Catalog window 2. Dataset RECOMMENDATION NITIFICATION button	NR	NA
3	User clicks on the Dataset "Request Recommendations" button	Click on "Request Recommendations" button	System displays: 1. Recommendation criteria 2. Email address field 3. Level of information required (dropdown menu) 4. OK button	NR	NA
4	User continues from step 3 of TC19.1-UC19 to the end	Tester continues from step 3 of TC19.1-UC19 to the end	Expected result continues from step 3 of TC19.1-UC19 to the end	NR	NA



Table 35: View charts

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC20.1-UC20	View charts	This test verifies the functionality which enables users to view charts of dataset on the platform	User must be a registered user be able to log on to the platform. User must choose the dataset to view	System perform the action by displaying a visualized view of the chosen dataset in a chart mode. Multiple chart tools should be available to user	Test data to be determined
Step #	Use case steps	Test Procedure	Expected Results (I = Internal)	Actual Result	Pass/Fail
1	User selects a dataset name on the dataset catalogue window to view in chart mode	Click on the dataset name on the dataset catalogue window	System displays the dataset on screen	NR	NA
2	User clicks on the VIEW CHART button	Click on the VIEW CHART button	System requests user to choose a determinant variable factor	NR	NA
3	User provides system requests	select variables from the from the dataset	system verifies the dataset (I)	NR	NA
4	None	None	System requests a web service (I)	NR	NA
5	None	None	system receives a chart (e.g. pie chart) from web service (I)	NR	NA
6	None	None	System display result on screen	NR	NA

Table 36: Login (to TET/SPOD-enabled CKAN platform)

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC23.1-UC23	Login (to TET/SPOD-enabled CKAN platform)	Verification of system capability to login a user into account on platform via TET or SPOD interfaces	1. User is a registered member of the platform 2. User remembers his valid login details. 3. System is connected to the internet and platform servers	1. User logs in to access resources 2. User is granted access to functions on system 3. User explore resources	Login details: UN = Tester; Password = testmaster1

Step #	Use case steps	Test Procedure	Expected Results (I = Internal)	Actual Result	Pass/Fail
1	User open platform window and browses the pages	Type the platform url on browser address bar	System open platform pages	NR	NA
2	User types in login details	Type in Login details: UN = Tester; Password = testmaster1	System validates entries and logs tester in system	NR	NA

Table 37: Login (Exceptional case - user enters invalid login details)

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC23.3-UC23	Login (Exceptional case - user enters invalid login details)	verification of system response to invalid login details in attempting to login on the TET/SPOD enabled CKAN platform	1. User is a registered member of the platform 2. User remembers his valid login details. 3. System is connected to the internet and platform servers	1. System should display warning message indicating use of invalid login Username or password 2. System should enable user retry the login attempt 3. System should login user when the correct details are used	Login details: UN = Tester; Password = testmaster2
Step #	Use case steps	Test Procedure	Expected Results (I = Internal)	Actual Result	Pass/Fail
1	User types in invalid login details	Type in Login details: UN = Tester; Password = testmaster2 and click on LOGIN button	System displays warning message: [Invalid username or password. Try again]; OK	NR	NA
2	User clicks on OK button and enters correct details then clicks on LOGIN button	Click OK button and retype UN = Tester; Password = testmaster1, then click on LOGIN button	System validates entries and logs tester in system	NR	NA

Table 38: Enrich profile

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC24.1-UC24	Enrich profile	This test verifies the functionality of the system which allows users to enrich a profile on the platform by adding some personal information such as occupation, interest or hobby, age, sex, religion etc.	User must be registered on platform and has access to his profile page. System tools must be available to enable user edit his profile status	System should seamlessly update user's profile when user is done with the changes. System should apply the updated user's status to treat user requests and activities on platform	Login details: UN = Tester; Password = testmaster1. After login, user change password to tester1, add hobby = Fishing, occupation = marine biologist
Step #	Use case steps	Test Procedure	Expected Results (I = Internal)	Actual Result	Pass/Fail
1	User open platform window and browses the pages	Type the platform url on browser address bar	System open platform pages	NR	NA
2	User types in login details	Type in Login details: UN = Tester; Password = testmaster1	System validates entries and logs tester in system	NR	NA
3	User clicks on ACCOUNT button	Click on ACCOUNT button	System opens personal profile page of user	NR	NA
4	User clicks on EDIT button	Click on EDIT button	System opens personal profile page of user	NR	NA
5	User types in all required changes and additions	Type in required additional fields and edit the required field	None	NR	NA
6	User clicks on DONE button	Click on DONE button	System updates all files and displays notification message [Profile successfully updated]	NR	NA

Table 39: Enrich profile (Alternative flow - user is not yet registered on platform but attempts to visit account settings page)

Test case ID	Test Case Name:	Description	Pre-conditions:	Post-conditions:	Test Data
TC24.2-UC24	Enrich profile (Alternative flow - user is not yet registered on platform but attempts to visit account settings page)	This test verifies the functionality of the system which prompts users to create account first on the platform before adding some personal information such as occupation, interest or hobby, age, sex, religion etc.	User must not be registered on platform but has access to browse platform pages that do not require login. System tools must be available to enable user create his profile status	System should block user's attempt to enter restricted pages. System should display warning message and option to register before visiting account page.	Login details: UN = Tester; Password = testmaster1. After login, user change password to tester1, add hobby = Fishing, occupation = marine biologist
Step #	Use case steps	Test Procedure	Expected Results (I = Internal)	Actual Result	Pass/Fail
1	User open platform window and browses the pages	Type the platform url on browser address bar	System open platform pages	NR	NA
2	User clicks on ACCOUNT button	Click on ACCOUNT button	System denies entry and issues warning message [You need to create an account first]; OK	NR	NA
3	User clicks on ACCOUNT button	Click on ACCOUNT button	System opens personal profile page of user	NR	NA
4	User clicks on OK button	Click on OK button	System opens Login form	NR	NA
5	User types in username and password in the required format and clicks on LOGIN button	Types in username and password in the required format and click on LOGIN button	System validates entries and create account for user	NR	NA
6	User continues from TC24.1, step 3 to step 6	Tester continues from TC24.1, step 3 to step 6	System results continues from TC24.1, step 3 to step 6	NR	NA

## BIBLIOGRAPHY

---

- [1] A Control Theory Perspective on Agile Methodology Use and Changing User Requirements, Likoebe M. Maruping, Viswanath Venkatesh, <http://pubsonline.informs.org/doi/pdf/10.1287/isre.1090.0238>
- [2] U-SCRUM: An Agile Methodology for Promoting Usability, Mona Singh, <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4599538>
- [3] The Scrum Guide™ The Definitive Guide to Scrum: The Rules of the Game  
Ken Schwaber and Jeff Sutherland, <http://www.scrumguides.org/docs/scrumguide/v1/scrum-guide-us.pdf>
- [4] Joomla. Drupal and WordPress - A Statistical Comparison of Open Source CMS. Savan K Patel  
, V.R. Rathod, Satyen Parikh. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6169111>
- [5] Quote from the CKAN 'README' file <https://github.com/okfn/ckan> (22/05/2013)
- [6] Open Data and the Academy: An Evaluation of CKAN for Research Data Management, Joss Winn, University of Lincoln, UK.1, <http://eprints.lincoln.ac.uk/9778/1/CKANEvaluation.pdf>
- [7] Beck, Kent; et al. (2001). *Manifesto for Agile Software Development*. Agile Alliance. Retrieved 14 June 2010.
- [8] Jen, L. and Lee, Y. (2000) IEEE recommended practice for architectural description of software-intensive systems, IEEE Architecture.
- [9] Tutorialspoint Online (n.d.) SDLC – The V-Model. [http://www.tutorialspoint.com/sdlc/sdlc\\_v\\_model.htm](http://www.tutorialspoint.com/sdlc/sdlc_v_model.htm)
- [10] Easterbrook Steve (2010). The Difference Between Verification and Validation. Serendipity blogpost. <http://www.easterbrook.ca/steve/2010/11/the-difference-between-verification-and-validation/> Viewed 09 Oct., 2015.
- [11] Tutorialpoint (Online) [http://www.tutorialspoint.com/sdlc/sdlc\\_v\\_model.htm](http://www.tutorialspoint.com/sdlc/sdlc_v_model.htm). Viewed 12 Oct., 2015
- [12] Vihari, Chaitanya (2014) Software Development Lifecycle Models. <http://www.testnbug.com/2014/12/software-development-life-cycle-models/>. Viewed: 13 Oct., 2015
- [13] Firesmith, Donald (2013) Using V Model for Testing. [https://insights.sei.cmu.edu/sei\\_blog/2013/11/using-v-models-for-testing.html](https://insights.sei.cmu.edu/sei_blog/2013/11/using-v-models-for-testing.html). Viewed 12 Oct., 2015
- [14] ISTQB (online) What is Agile Model – Advantages and Disadvantages. <http://istqbexamcertification.com/what-is-agile-model-advantages-disadvantages-and-when-to-use-it/>. Viewed 12 Oct., 2015
- [15] Institute of Electrical and Electronics Engineers. (IEEE 829) Software Testing Template.
- [16] Eric Bissonnette 2015. Test Plan Part 1: Why Create a Test Plan. Source: <https://crowdsourcedtesting.com/resources/why-create-a-test-plan/> Viewed: 07 Oct., 2015
- [17] ISTQB (n.d) What is the Purpose and Importance of Test Plans in Software Testing? <http://istqbexamcertification.com/what-is-the-purpose-and-importance-of-test-plans/> viewed 07 Oct., 2015
- [18] Purpose & Advantages of Test Plan (n.d) <http://www.softwaretestingtimes.com/2010/05/purpose-advantages-of-test-plan.html>
- [19] Easterbrook Steve, Professor, Department of Computer Science, University of Toronto, Canada.