



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



Project number 645860
H2020-INSO-2014

D5.1 Research Report Year 1

(01/30/2016)



WISE&MUNRO



Document produced by

Organisation: Wise & Munro

Author / email: Jerry Andriessen / jerry@wisemunro.eu

Subject: Research Report

Due date: Jan 31, 2016

Dissemination level: PU**Reviewed and approved by**

Date	Name	Organization
Jan 28	Francoise Detienne	CNRS

Revision History

Version	Date	Authors	Status	Description of Changes
0.1	11-06-2015	JA, MP	Draft Outline	Concept
0.2	16-12-2015	JA, MP	Alpha	New structure, addition of content
		MB, FD, JF		Contribution CNRS Paris: activities and community research approach
	20-12-2015	ER, AM		Contribution UU: society model and general evaluation model
0.3	14-01-2016	JA		Thorough Revision
	20-01-2016	VS		Contribution Salerno regarding SPOD
	22-01-2016	AO, MH		Contribution Galway: User level model and research
0.4	22-01-2016	JA, MP		Contributions from Prato, Salerno, UU, Galway and Paris further integrated
0.5	25-01-2016	JA, MP	Alpha	Integration of Evaluation Model, version for revision
1.0	30-01-2016	JA, MP	Final draft	Integration of Reviewers' comments

EXECUTIVE SUMMARY

The main goal of the Route to PA project is to achieve greater transparency of local government, through the organisation of a number of activities and scenarios with Open Data, mediated by our technology. Five local governments participate in providing cases for achieving this goal within their contexts: Prato, Groningen, Dublin, Issy-les-Moulineaux, and Den Haag. A year of negotiations and preparations has passed. These are summarised in this Deliverable. This Work package is about our approach for evaluation of the tool, especially concerning increased transparency of local government.

We first briefly introduce the five settings where we work on introducing our tool (Chapter 1). Our descriptive approach can be characterised as chronological:

1. To identify user needs and tool requirements, leading to evaluation requirements at the level of interaction with our system (WP2, Chapter 2).
2. To identify the organisational situation in terms of a societal activity model, thereby identifying organisational objectives and tensions at the five sites, and their relations and expectations concerning citizens supposed to work with the tool (Task 3.1, Chapter 3).
3. To identify the specific characteristics and objectives of groups of users actually working with the tool, and to study their activities with the tool in terms of transparency-enhancing scenarios (T3.2, Ch. 4).
4. To report on activities at the user sites in relation to upcoming pilot studies (Chapter 5).
5. To translate these levels into a general approach to evaluation at all sites (Chapter 6).
6. To develop research plans for evaluation at the five sites (Chapter 7).

This means that for **evaluation** of the effects of our approaches at the five user sites, we have several sources of criteria for evaluation. The identification of users' needs translated into design specifications for the tool, that is, the actions users are supposed to perform with the tool (WP2) leads to a set of evaluation criteria **at the technology- and user levels** (Chapter 2). The **modelling** activity (WP3) has provided abstract models at the **society** and **community**-levels. On the basis of these models we will be able to characterise and compare all cases with respect to their current and future states at two levels: that of the **community**, or small group working together to generate new ideas, consisting of users carrying out different actions within the context of some objective or assignment (Chapter 4), and that of the **society**, or the organisational context, adapting to the new possibilities for creating transparency and facilitating use of open data in the organisation (Chapter 3). Chapter 5 will provide a summary of ongoing work at the **user sites**.

These research plans have a general component, and a site specific component. In the general component, we elaborate the four levels of description of activity (see above: technology, users, community, and society), that all will be studied next year to provide a serious answer to the questions about the roles of the tools to enhance transparency through open data in local government. Our project is an implementation project, not a research project, so we mainly focus on this main question and the main outcomes at the four levels. As a design based approach, every series of evaluation activity results in new recommendations at the four levels of description of the activities. The first cycle of such evaluation starts at the beginning of March 2016 (M14). In Chapter 7 we elaborate our (tentative) site-specific **research-plans** for the upcoming year, and more concretely for the period M13-M18.

It should be underlined that our efforts aim at, on the one hand, respecting the views and state of development of open data at the five sites, and, on the other hand, harmonising our evaluation approaches as much as we can. Most planning is ongoing, and details of it depend on development of the affordances of the tools in use at each site after February 2016. After delivery of the alpha versions of both TED and SPOD, we will prepare further evaluation plans for (the first part of) year two. These will be presented in D5.2 (M24).

TABLE OF CONTENTS

1	Introduction	5
1.1	Months 1-6	5
1.2	Months 6-12	7
2	User- and technology levels of evaluation	9
2.1	Summary of WP2 work on User Needs	9
2.2	Summary of system requirements	12
2.3	The technology level	17
3	The societal level of activity	20
4	The community level.....	21
4.1	Community building	21
4.2	Features of collaboration	24
4.2.1	Collaboration.....	24
4.2.2	Phases during the working process.....	25
4.2.3	Discussion styles	26
4.2.4	Use of open data during discussions.....	27
4.2.5	Moderation	28
4.3	How route-to-pa facilitates collaboration	29
5	Activities at the user sites M6-m12	30
5.1.1	Prato.....	30
5.1.2	Groningen	32
5.1.3	Dublin.....	33
5.1.4	Cnrs/Issy-les-moulineaux	35
5.1.5	Wise&munro/Den haag	37
6	Evaluation and transparency	40
6.1	Views on transparency	40
6.2	Scenario for the evaluation plan for year 2	41
6.2.1	Object of Evaluation: Activity Model	41
6.2.2	Functions of Evaluation: Assessment, Design and Local Development	42
6.3	Levels of evaluation: general plans	43
6.3.1	The society level.....	43
6.3.2	The community level.....	44
6.3.3	Analysis at the user level.....	49
7	Research plans for year 2	52
7.1	Prato	52
7.2	Groningen.....	54
7.3	Dublin	55
7.4	CNRS Paris	56
7.5	Den Haag/Wise & Munro	58

1 INTRODUCTION

The main goal of the Route to PA project is to achieve greater transparency of local government, through the organisation of a number of activities and scenarios with Open Data, mediated by our technology. Five local governments participate in providing cases for achieving this goal within their contexts. A year of negotiations and preparations has passed. This Work package is about evaluation of the tool and our approach, especially concerning increased transparency of local government, during the first year (of three).

As a consequence of cultural variations, local interpretations of transparency and tool requirements for the cases are different. Such a situation will lead to rich data, but also to more complicated models needed for comparison and recommendations. Therefore, although we try to achieve a level of coherence between the cases concerning the timing and the applied methods of evaluations (see section 4.3), complete equivocality will not be possible. As a consequence, some section will be quite particular to one specific case, but other cases will pick up only those aspects (and as much of those as possible) that they can use for their own evaluation process.

1.1 MONTHS 1-6

The first six months of the project have been used for setting up groups of participants and stakeholders, discussing their general issues (barriers) concerning communication between local government and various stakeholders and the various possibilities for open data to support overcoming these barriers in collaborative or instructive dialogues between public administrations and citizens (D2.1, 2.2, 2.3, 2.4).

The Route-to-Pa project currently distinguishes 5 **cases** (text based on D3.1 descriptions):

- 1) The Prato scenario case is about city budget management. The goal is for citizens to monitor the allocation of the city budget by municipality and to propose expenditure priorities and suggestions. The administration, both at political and bureaucratic levels, aims to increase transparency on budget management by enhancing dialogue with citizens by soliciting citizens' expenditure suggestions. The administration has made a commitment to open data. In its open data policy document (Prato, 2014), the municipality of Prato emphasizes that the published datasets are owned by the community, and citizens have the right to access and freely reuse the open data (Comune di Prato, 2014, p. 3). The city would like to increase transparency on budget management and possibly collect some of citizens' expenditure suggestions. It is important to know that this process is intended to be an improved version of existing democratic procedures of involving citizens in budget decisions.
- 2) Groningen is the most north-eastern Province of the Netherlands. The capital of the province Groningen is the city of Groningen with 189.000 inhabitants. The Province and the City of Groningen have their joint open data portal (www.data.groningen.nl) with more than 70 datasets and visualizations of financial data and health concerns. The object or policy issue in Groningen focuses on population decline. One of the areas with the highest population decline in the Netherlands is situated in Groningen. The societal issue of population decline affects several policy areas e.g. health care, education, employment, housing and services. The pilot will focus on the issues of housing,

employment and healthcare, in line with the coalition agreement. The main policy aim is to guide and to anticipate the consequences of population decline. The administration facilitates citizen participation but would like to increasingly collaborate or co-produce with citizens among others by stimulating citizens' initiatives. The potential community of stakeholders is diverse consisting of citizens, public organizations (schools, health care organizations etc.) and private companies. Open data as an instrument can provide insight in the consequences of and might be able to contribute to innovative and collaborative solutions for population decline.

- 3) Dublin, Ireland ("Dublinked"): The Dublin City Council is one of the founding partners of Dublinked. Dublinked is an ideas and information sharing network which connects the Dublin region's four local authorities with universities, companies and entrepreneurs. Dublinked brings people together to test new ideas using live city data and to develop new products and services using the city as a testing ground (Dublinked, 2015). The objective of Dublinked is to enable innovative applications by entrepreneurs and businesses in areas like public transportation, planning, social services and public facilities (Hogan, et al., 2015, p. 19). In accordance with the Digital Masterplan of Dublin: "Open Data is a key element of the commonage and the initiative taken regionally in Dublinked should be extended and developed" (Dublin Digital, 2013, p. 13). The scenario chosen for ROUTE-TO-PA evolves around capacity building, which focuses on increasing citizen engagement in a deliberative process with their city or more specifically their community. The Dublin City Council for instance launched the program Your Dublin, Your Voice, an online feedback survey that "gives Dublin's citizens and visitors a unique opportunity to provide opinions and views on what they love, like and dislike in our city." (Dublin City Council, 2015). The scenario thus focuses on building community awareness with the ultimate outcome to make my city great.
- 4) Issy- les- Moulineaux, France (young entrepreneurs): Issy-les-Moulineaux is a city located near Paris. The city hosts many IT companies (Issy-les-Moulineaux, 2015) and welcomes start-up companies in the field of new technologies. The CNRS-Issy-les-Moulineaux scenario focuses now on the activity of two central groups: young entrepreneurs in ICT domain who form a major part of the local economic base and public administrators from Paris Region who collaborate with Issy-les-Moulineaux in the global open data policy. So, the CNRS team is interested on (i) young entrepreneurs/users' needs in terms of data accessibility and usability, (ii) PAs/producers' expectations concerning the economic exploitation of open data and (iii) the possible tensions between these two activity systems (young entrepreneurs versus PAs).
- 5) The Hague, the Netherlands (employers and employment): The pilot in The Hague focuses on collaboration between public administrators and employers and can be characterized as a participatory process. Employers and the City of The Hague have a longer history of collaboration and meeting, the relationship between the Department and the local employers is quite good. The proposed solution will help them to achieve better solutions for existing labour market issues. The specific scenario or policy issue suitable for exploiting Open Data will be jointly developed, whereby the focus is on finding solutions for existing problems together. This is called co-creation. Possible topics could be developing innovative solutions for the increasing number of jobless elderly people or developing practical implications for new legislation about employing difficult people such as handicapped individuals or long-term jobless.

It can be said that, at a general level of description, all cases involve some degree of community building for developing new products and services. The strategy to follow depends on the goals and natures of the different communities, including their readiness to exploit community services and open data. In two of the cases (1 and

3), the community building is a consequence of large groups of individual citizens (or citizens involved in similar interactions with their municipality) using the tools and developing a sense of togetherness as a consequence of their (individual or collective) engagements with their municipality. We call this type of use **consultation**. In the other three cases (2, 4, and 5), constructed small groups of mixed types of citizens collaborate mediated by our tools and share their products with their already existing communities, with the possibility of further dissemination of the tool and the approach as useful for dealing with similar issues within the community, or between the community and the municipality. This type of use can be characterised as **collaboration**. The goal for using open data in both types of cases is to create new, principled and better solutions for existing community-based issues, and, as a consequence of this process transparency will be increased: a broader basis for consultation, participate in decision-making, and understanding mutual constraints and possibilities.

1.2 MONTHS 6-12

During the second half of the first year we focused on two things: (1) developing the first prototypes of SPOD and TET, according to the specifications of WP2 and the outcomes of several iterations of usability work (see D4.1) until the alpha version will be produced in M13; (2) engaging users and their sites in preparatory activities, further characterising and comparing the sites in society-models, and setting up research plans and activities for period M13-M20.

During the first phase of technology developments, the main developments from M6 until month 12 have focused on the design specifications developed in the user needs activities (D2.4), and less on the more specific affordances of a collaboration space. A workshop planned for M13 will focus on further specifications for collaborative use, e.g. structuring dialogues, group spaces for particular collaboration types (see chapter 4), and designing for multiple visualisations being visible. There also is a parallel division in preparatory activities between the collaborative and consultation cases, with the last putting more stress on the internal technical and support organisation onsite, while the first have more attention for particular collaborative scenarios to be played out with the tools in small groups. It should also be said that current differences between sites relate to local differences in approach to implementation, related to current states of working with open data, rather than on fundamentally different views on research within the project. As an implementation project, we think it is crucial to adapt our research approach to objectives in the local contexts.

The **user needs** (WP2) have provided us with a rich source of information about what our users want. In addition, theory- and evaluation criteria for verification and validation of technology can be directly derived from the WP2 work. Chapter 2 provides a summary of user needs work and how this leads to the verification and validation work at all sites. We also provide a summary of the tool, highlighting some of the criteria developed in WP2. In Chapter 4 we provide a short theoretical section which discusses a couple of critical issues concerning collaboration with social media, on the basis of open data. On the basis of our knowledge and experience two recommendations follow: 1) to engage potential users in a preparation phase, to familiarise them with the social actions that are required from them; 2) to design specific evaluation criteria for each user site to capture the essential features of each case with respect to working with open data. Ongoing preparation is discussed in chapter 5. As we said, two types of scenarios have developed, having different angles at community building. The **consultation** type has the explicit goal of users of the technology becoming skilled at building and sustaining of communities of participants (Lave & Wenger, 1991; Meijer, Grimmelikhuijsen &

Brandsma, 2012). The **collaboration** view on communities focuses on smaller collaborative groups discussing some issue, supported by datasets and further tool affordances. Such small groups need a closed space with easy overview of the functionalities and activities focused on open data. The **consultation** types of scenario are directed at creating more transparency through **community-building**, the second type at creating more transparency through **collaborative discussions** and sharing its outcomes. Although this looks like a strong contrast, evaluation of the success of community building involves looking at collaborative discussions, whilst evaluation of the success of collaborative discussions also requires looking at the impact on the community. Most probably, several scenarios are useful at all sites, and our initial distinction may turn out to be two aspects both required for open data social sites.

Concerning **evaluation** of the effects of our approaches at the five user sites, we have several sources of criteria for evaluation. The identification of users' needs, directly translated into design specifications for the tool, that is, the actions users are supposed to perform with the tool (WP2) leads to a set of evaluation criteria **at the technology and user levels**. The **modelling** activity (WP3) has provided abstract models at the **society** and **community**-levels. On the basis of these models we will be able to characterise and compare all cases with respect to their current and future states at different levels: that of the technology functioning according to design specifications, that of the individual user carrying out different actions, that of the community, or small group working together to generate new ideas, and that of the society, or the organisational context, adapting to the new possibilities for creating transparency. Chapters 3 and 4 will provide a summary of this work, and in chapter 7 we elaborate our **research-plans** for the upcoming year, and more concretely for the period M13-M18. These research plans have a general component, and a site specific component. In the general component, we distinguish four levels of description of activity (technology, users, community, and society), that all will be studied next year to provide a serious answer to the general questions about the effectivity of the tools to enhance transparency through open data in local government. Our project is an implementation project, not a research project, so we focus on this main question and the main outcomes at the four levels. As a design based approach, every series of evaluation activity results in new recommendations at the four levels of description of the activities. The first cycle of such evaluation starts at the beginning of March 2016 (M14). The **technology** construction activity (WP4) has been working on the construction of a tool to support users in finding, analysing, arranging, visualising and discussing open data. At this moment, the first prototypes of our tools will be ready for a first round of testing by our user groups. Reports on usability work at the technology level are in D4.1.

2 USER- AND TECHNOLOGY LEVELS OF EVALUATION

As we have explained in the introduction, WP5 builds further on work in WP2 on user needs and system requirements. This work leads to an approach to evaluation at this level.

2.1 SUMMARY OF WP2 WORK ON USER NEEDS

A series of carefully designed workshops were conducted, one in each pilot site, for the purpose of developing a comprehensive set of user needs, as proposed by key stakeholders. Each workshop brought together experts, academics, industry specialists, open data practitioners, representatives of governments, open data researchers, and potential users (including citizens, representatives of citizens and social service institutes, and journalists) to brainstorm on open data platform adoption challenges, solutions to the challenges and a set of needs and requirements necessary for consideration in the design of the ROUTE-TO-PA platform. The emphasis on citizen participation and collaborative design in the methodology seeks to address the goals of improved government transparency and accountability for decision-making. Each workshop began with a collective intelligence (CI) analysis of barriers to accessing, understanding and using open data, followed by an analysis of options that may overcome these barriers. Participants then worked to develop scenario-based user needs, which involved profiling user needs in light of the barriers and options and high level scenarios of open data usage.

In a typical CI session, a group of participants who are knowledgeable about a particular situation engage in (a) developing an understanding of the situation they face, (b) establishing a collective basis for thinking about their future, and (c) producing a framework for effective action. In the process of moving through these phases, group members can develop a greater sense of teamwork and gain new communication and information-processing skills. CI utilizes a carefully selected set of methodologies, matched to the phase of group interaction and the requirements of the situation. The most common methodologies are the nominal group technique, ideawriting, interpretive structural modeling, and field and profile representations. For the purposes of idea generation in this context, the ideawriting technique was used, along with categorisation or field representation of ideas.

Ideawriting is a method that utilizes relatively small groups of 4-6 persons each, formed by dividing a larger group into several working teams, for the purpose of developing ideas and exploring the meaning of those ideas through open discussion. Ideawriting involves five steps: (a) presentation of a stimulus question to participants; (b) silent generation of ideas in writing by each participant working alone; (c) exchange of written sheets of ideas among all group members, with opportunity for individuals to add ideas as they read others' papers; (e) discussion and clarification of unique ideas; and (f) an oral report of the ideas generated by each working group in a plenary session. In this plenary session, duplicate ideas across the working groups are eliminated from the set and new ideas (if any) are added; the resulting set of ideas is then ready for use in the next stage of the group's work.

In the current application of CI, workshop participants worked to develop scenario-based user needs, which involved profiling user needs in light of the barriers and options and high level scenarios of open data usage. This included a separate focus on (1) information needs, (2) social and collaborative interaction needs, and (3) understandability, usability and decision-making needs. Idea writing was used for each cluster of needs. High

level scenarios including multiple users were used to prompt thinking in relation to user needs. All the short user stories generated by participants were generated in the form:

As User Type _____, I want _____, so that I can _____

After addressing each set of needs for the first set of scenarios, a second set of user scenarios was introduced for discussion and idea generation. The wants (or needs) generated by participants across each pilot site were then analysed and key categories of user needs identified. Reasons for specified user needs were also analysed, and this analysis was used to advance our understanding of the scenarios and prospective use case models. A major innovation in this project is the combination of collective intelligence methodologies (Warfield, 2006) with scenario-based design (Carroll, 2000) and agile user story (Cohn, 2004) methods in the analysis of needs (D2.3) and the generation of use case models and system requirement (D2.4). This work in turn has shaped the test and evaluation framework (see Figure 1).

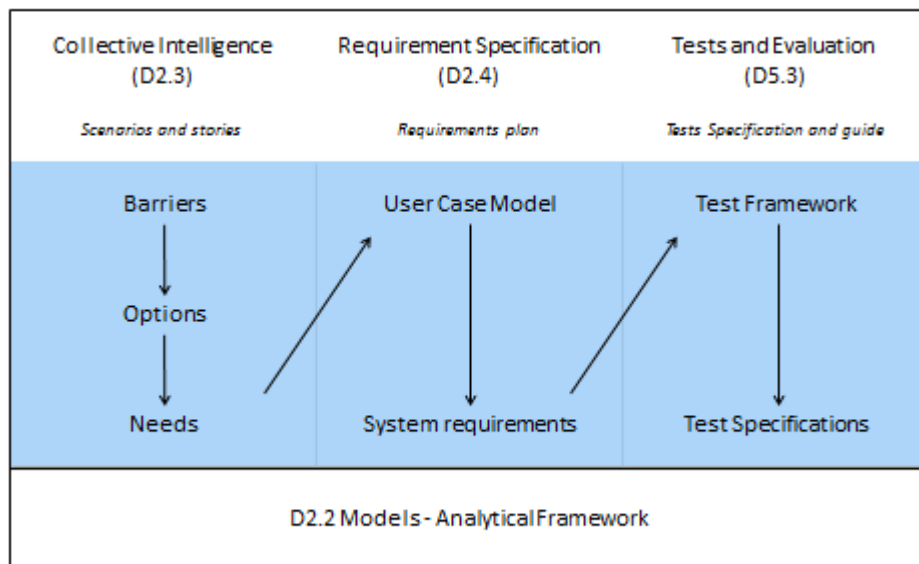


Figure 1. Workflow for WP2 and link between WP2 and WP5 evaluation framework and test specifications

Central to our work is a scenario-based design approach. Broadly speaking, the scenario-based design framework describes an iterative approach to interactive systems design and analysis, and encourages a reasoning process about people using technology and about finding trade-offs throughout development, including trade-offs between the potential impact of design decisions and the feasibility of the design options. One specific aim of the CI workshops was to elicit and document user needs from the scenarios provided. These scenarios, which were consistent with the pilot case in each site, were designed to facilitate discussion over the needs and requirements of open data users. Scenario-based design is a set of techniques in which the use of a future system is described at an early point in the development process. Narrative descriptions of envisioned usage episodes help to guide the development of the system that will enable these user experiences. Importantly, scenarios are work-oriented design objects. They describe systems in terms of the work that users will try to do when they use those systems. In this way, they help to ensure that design work remains focused

on the needs and concerns of users. Scenarios can be made even more effective as work-oriented design objects when users are directly involved in creating them. Ackoff (1979) noted that the indeterminacy of design situations makes it imperative that all stakeholders participate directly. This is central to the approach taken in WP2, where we built upon the collective intelligence of stakeholders in facilitating development of scenarios and user stories. Importantly for ROUTE-TO-PA, scenario-based design and both the process and products of this work supports a process of participatory design, where prospective users begin by enacting or relating episodes of current activities, then work iteratively with designers to transform and enrich these scenarios with the opportunities provided by new technologies.

The scenarios used in WP2 addressed various contextual issues, relevant to each of the workshop sites, and aligned with the primary case focus in each pilot site. For example, the Dublin workshop focused on *community networking and opportunity creation*; the Groningen workshop focused on the use of Open Data in overcoming issues associated with *population decline*; the Den Haag workshop focused on Open Data in relation to *employment and opportunity creation*; the Prato workshop centered on *local policy and budget issues*; and finally, the workshop in Paris focused on Open Data in relation to *start-up companies and the digital economy*. As such, there was some variety in user needs generated, across all there categories of needs: information needs, social and collaborative needs, and understandability, usability and decision-making needs.

As workshop participants in each pilot site were working with a variety of scenarios, the user information needs generated were numerous and diverse. The information needs included, for example, demographic information needs; legal information needs; health information needs; social and community information needs; planning information needs; services, amenities and event information needs; business and financial information needs; jobseeker information needs. Essentially, the data and information that different pilot sites need depend on the problems they are working to solve in their scenarios. The ROUTE-TO-PA team are working to collate all available open data to make it available on the platform.

Broadly speaking, the scenario-based design framework describes an iterative approach to interactive systems design and analysis, and encourages a reasoning process about people using technology and about finding trade-offs throughout development, including trade-offs between the potential impact of design decisions and the feasibility of the design options.

One specific aim of the CI workshops was to elicit and document user needs from the scenarios provided. These scenarios, which were consistent with the pilot case in each site, were designed to facilitate discussion over the needs and requirements of open data users. Scenario-based design is a set of techniques in which the *use* of a future system is described at an early point in the development process. Narrative descriptions of envisioned usage episodes help to guide the development of the system that will enable these user experiences.

Subsequent to gathering and integrating scenario-based user needs across all pilot sites, the ROUTE-TO-PA design team engaged in an exercise designed to rate the relative impact and feasibility of specified needs. This resulted in the first set of user needs selected for agile software development from M6 to M12, and this process continues iteratively into Year 2 as the design team revisits user needs and ways in which SPOD and TET design features can support those needs.

The information needs included, for example, demographic information needs; legal information needs; health information needs; social and community information needs; planning information needs; services, amenities and event information needs; business and financial information needs; jobseeker information needs. Essentially, the data and information that different pilot sites need depend on the problems they are working to solve in their scenarios. The ROUTE-TO-PA team are working to collate all available open data to make it available on the platform.

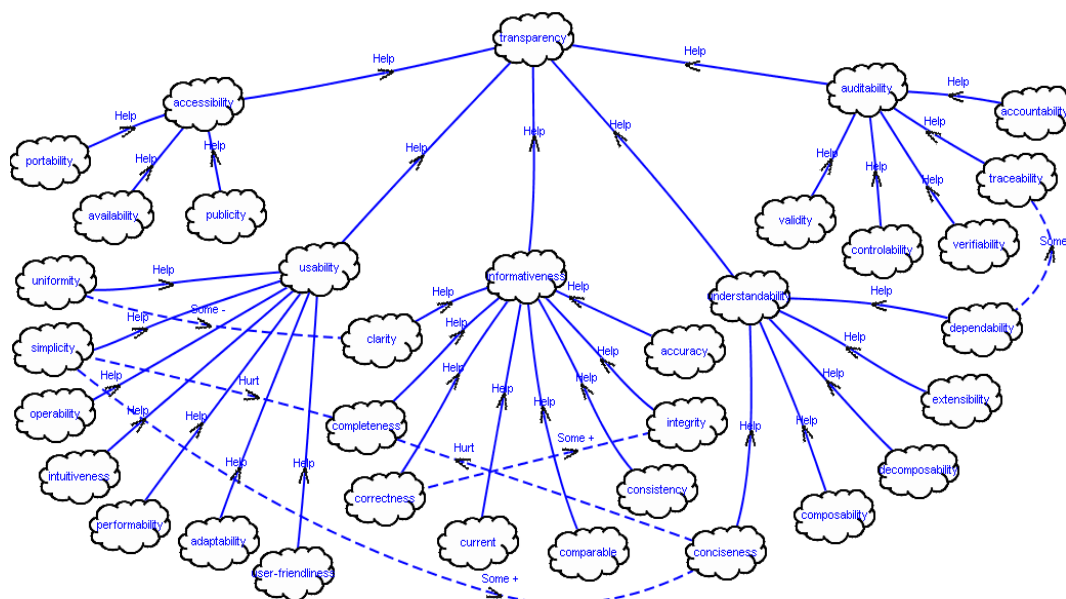
Social and collaborative needs were commonly specified across pilot sites and included dialogue and discussion spaces; moderation and maintenance of these spaces; platform tool capabilities for interaction; varied forms of social media interaction; personalisation of user spaces; and requesting and sharing information. Broadly speaking, participants identified a variety of forms of interaction which could be used over Open Data, and suggested a number of considerations and affordances which would increase the impact and appeal of such social and collaborative platforms (see D2.1 for more details).

Similarly, the range of understandability, usability and decision-making needs across sites commonly included: Affordances for the visualisation of complex information; data analysis and reporting tools; decision-making support tools; guidance and usage support tools; affordances for personalising platforms and/or data; and certification tools. Broadly speaking, participants frequently cited the need for data visualisation tools, among others, which would make data more easily understood, whether for personal or professional use.

2.2 SUMMARY OF SYSTEM REQUIREMENTS

The user needs, as determined by key stakeholders at the CI workshops across pilot sites and cases provide a comprehensive and detailed insight into what key stakeholders want from the platform. As such, it is possible to derive criteria from this information for the purpose of evaluating user experience, including validation of the tools in line with user needs. As these user needs were based on the wants and needs of users in response to open data usage scenarios, any instrument designed to evaluate the usability and capabilities of the tools should incorporate the criteria as determined by stakeholders in the CI sessions.

The CI work which was conducted with key stakeholders has provided an extensive and powerful set of user needs and requirements across three high-level categories: (1) Information needs, (2) Social and collaborative needs, and (3) Understandability, Usability, and Decision-making needs. It is our view that these three high-level categories can serve as a focus of evaluation in the context of the broader transparency components in the model below (see Figure 2). The SPOD and TET tools should seek to address these needs as a means to enhance the accessibility, usability, informativeness, understandability, and auditability of ROUTE-TO-PA platform, in accordance with user requirements. Section 2.3 explains how these features are developed in the tools.



More generally, evaluation work takes place across societal (section 6.2), community (section 6.3) and user (section 6.4) levels, and for each of these levels the research methodologies adopted will allow for a different lens of enquiry in relation to evaluating key information, social and collaborative, and decision-making needs (see figure 3). Notably, we have proposed a multi-level model that provides a cross-cutting evaluation of the extent to which ROUTE-TO-PA supports the range of information, social and collaborative, and decision-making needs identified by stakeholders in D2.3.

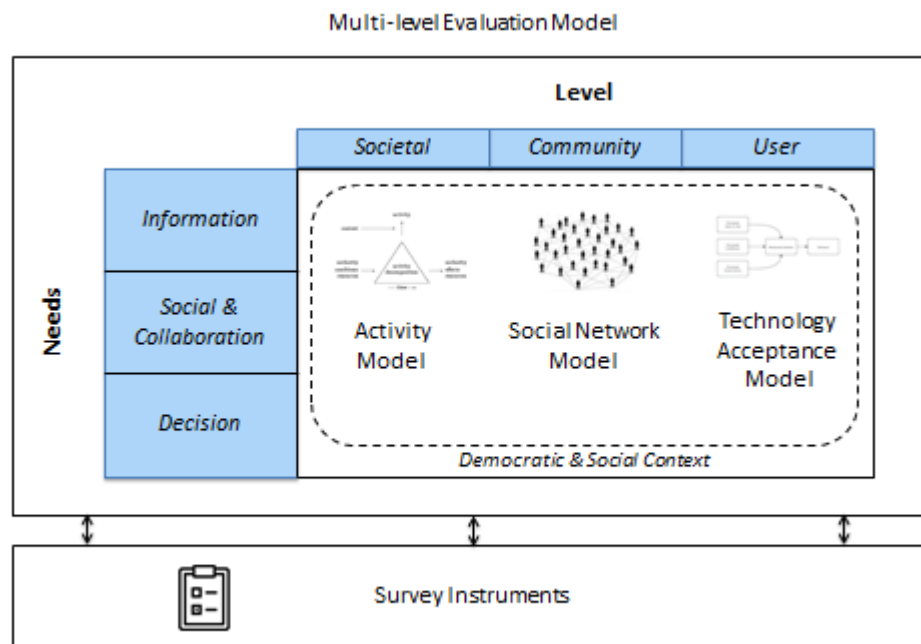


Figure 3. Multi-level model for evaluation of information, social and collaborative, and decision-making needs across societal, community, and user levels.

As noted above in section 2.2, M6 – M12 have involved intensive activity in translating user needs into key features that support those needs (see Tables 1, 2, and 3) and this work will continue iteratively into year 2 (see Tables 4 and 5). The core user groups in each pilot site will continue to provide collective intelligence input on the features available in successive iteration of the Route-to-PA platform, as part of a cyclical validation and evaluation process. The core methodologies involved at different levels of analysis, including the societal activity modelling for the societal level, the social network analysis for the community level, and the technology acceptance for the user level will provide a framework for iterative evaluation of the Route-to-PA platform. As illustrated below, each evaluation cycle will feedback into the iterative design work and the goal is for the platform to increasingly, over successive design cycles, be evaluated by stakeholders as increasingly valid in promoting user needs.

Table 1: Open data sources and datasets currently supporting information needs

Demographic information needs	<ul style="list-style-type: none"> Ancitel CKAN (http://ckan.ancitel.it/dataset) has datasets in Italian language about the municipalities, territorial and demographic information over the years Columby (https://www.columby.com/) has datasets about population in Den Haag
Legal information needs	
Social and community information needs	<ul style="list-style-type: none"> Issy (https://data.issy.com/page/home/) has datasets on public transports, Culture, Restaurants. Etc. OpenData CBS (http://opendata.cbs.nl/dataportaal/portal.html?_la=nl&_catalog=CBS)
Planning information needs	<ul style="list-style-type: none"> OpenData Gov (https://data.gov.uk/) OpenData CBS (http://opendata.cbs.nl/dataportaal/portal.html?_la=nl&_catalog=CBS)
Services, amenities and event information needs	<ul style="list-style-type: none"> Issy (https://data.issy.com/page/home/) https://data.gov.ie/data
Business and financial information needs	<ul style="list-style-type: none"> Issy (https://data.issy.com/page/home/) OpenBilanci (http://www.openbilanci.it/) Open Data Network (http://www.opendatanetwork.it/) Gronongen OpenData (http://data.groningen.nl/) (http://denhaag.buurtmonitor.nl/) https://data.overheid.nl/

Table 2: Features implemented to address users' Social and Collaborative Needs

Dialogue and discussion space	<ul style="list-style-type: none"> In SPOD the Agora enables the public discussion around Open Data visualisations, allowing the creation of public room with a topic and the posting of content and data visualisations (<i>Deliverable 4.1 Requirements 19-31 – public room</i>) Chat (<i>Deliverable 4.1 Requirement 6 – chat</i>) Messages (<i>Deliverable 4.1 Requirement 5 - private messages</i>)
Feedback	<ul style="list-style-type: none"> Any user can provide feedback by sending emails using the links shown on the SPOD platform. (<i>Deliverable 4.1 Requirement 46-49 - support to users</i>)
Forms of interaction	<ul style="list-style-type: none"> Discussion around visualisations of open data. (<i>Deliverable 4.1 Requirement 19-31</i>) Sentiment indicator for each post reply. (<i>Deliverable 4.1 Requirement 19-31 - public room</i>)
Live viewing	<ul style="list-style-type: none"> Open Data visualisation in real-time client side: Datalets (<i>Deliverable 4.1 Requirement 44-45</i>)

Local issues	<ul style="list-style-type: none"> Discussion of local issues using the public rooms within the Agora (e.g., discuss local issues) (<i>Deliverable 4.1 Requirement 19-31</i>)
Moderation and maintenance	
Personalisation	<ul style="list-style-type: none"> Private room where the users can privately store, digest and retrieve information (e.g., Text content, datasets visualisations, links to external resources) (<i>Deliverable 4.1 Requirement 41-43 – private room</i>)
Platform tool capabilities for interaction	<ul style="list-style-type: none"> Agora with public room (<i>Deliverable 4.1 Requirement 19-31 - private messages</i>) What's new: advertisement of public rooms activities to engage all platform users (<i>Deliverable 4.1 Requirement 17-18 – What's new</i>) My friends Activities shows the activities of the friends (<i>Deliverable 4.1 Requirement 19-31 - private messages</i>)
Sharing and requesting protocols	<ul style="list-style-type: none"> The creation and posting of data visualisations and sharing (<i>Deliverable 4.1 Requirement 32-40 - dataset visualisation</i>)
Standardised protocols	

Table 3: Features implemented to address users' Understandability, Usability, and Decision-Making Needs

Affordances for the visualisation of complex information	Chart for Dataset Summaries - TET Provides Charts to summarize dataset contents based on the aspects (or Columns) of the dataset that are of interest to the users.
Data analysis and reporting tools	In addition to the standard data exploration features available the base CKAN platform, TET provides features to summarize datasets using Pivot Tables. Pivot tables allows grouping of data based on one or more columns in the dataset. Charts can be automatically generated based on these summaries. Not supported at present.
Decision-making support tools	
Guidance and usage support tools	Online help features are currently available to guide users when needed. Videos on use of platforms will be provided later.
Affordances for personalising platforms and/or data	Two types of personalization features are provided by TET. The first enables available datasets to be categorised based on type or category of users in addition to related sectors. The second enables filtering of dataset search results based on both user categories and their preferences.
Certification tools	Not supported at present.

Table 4: Year 2, Features planned to address users' Understandability, Usability, and Decision-Making Needs

Affordances for the visualisation of complex information	TET will provide Chart Templates for one-click visualizing different categories of datasets (e.g. time series data, textual data). In addition infographics will be developed to support datasets visualisation with high transparency value (e.g. Budget balances data and data on facilities available in specific areas of a city).
Data analysis and reporting tools	TET will provide a simple Data Story templates for related datasets in addition to the Pivot Tables currently available on the platform. For advanced users, TET will enable users to integrate related datasets, develop queries to extract needed information and carry out explanatory analytics to check simple hypotheses about the datasets.
Decision-making support tools	TET will enable the construction of workflows comprising a sequence of simple analysis steps to produce necessary information to support a decision needs of users.
Guidance and usage support tools	Videos to guide users on the use of different aspects of the system will be developed and directly accessible through the TET platform.
Affordances for personalising platforms and/or data	In addition to the two types of personalisation provided already implemented in Year 1, TET personalisation will use information about dataset usage to present and recommend datasets to users.

Table 5: Year 2, Features planned to address users' Social and Collaborative Needs

Dialogue and discussion space	<ul style="list-style-type: none"> Threaded chat along with the collaborative space.
Feedback	
Forms of interaction	<ul style="list-style-type: none"> Collaborative space: will allow the creation and discussion of content among small groups of users Events
Live viewing	<ul style="list-style-type: none"> Integration not only form datasets but also for query able databases Most used visualisations Statics about datasets, visualisations and user generated content (rooms, text, ...)
Local issues	<ul style="list-style-type: none"> Collaborative Space: the content created from small groups of users with also an expert user could be public available to the community
Moderation and maintenance	<ul style="list-style-type: none"> Introduction of the user roles with specific moderator role.
Personalisation	<ul style="list-style-type: none"> Personalised search of datasets and content based on the user profile information. Suggested datasets and datasets visualisations

Platform tool capabilities for interaction	<ul style="list-style-type: none"> • Collaborative space. • Open data access: Advanced support for charts
Sharing and requesting protocols	<ul style="list-style-type: none"> • Events
Standardised protocols	

2.3 THE TECHNOLOGY LEVEL

We are designing, implementing and deploying a sustainable infrastructure for Open Data and effective citizen engagement strategies, in order to support users in finding and using in the best way open data. TO this purpose, we are working on two tools, a Social Platform of Open Data (SPOD) and a Transparency-Enhancing Toolset (TET) that are jointly designed in order to offer to citizens advanced services.

First, SPOD is a social platform designed to enhance and make transparency the value of data between the Public Administration and citizen through interactive visualization of dataset made available from Open Data portal powered by the TET Platform (compatible also with any other CKAN portal and with OpenDataSoft). SPOD provides an user-friendly interface to publish and start social discussion exploring datasets of the Open Data portals. SPOD is available in different languages (English, Italian French, Dutch).

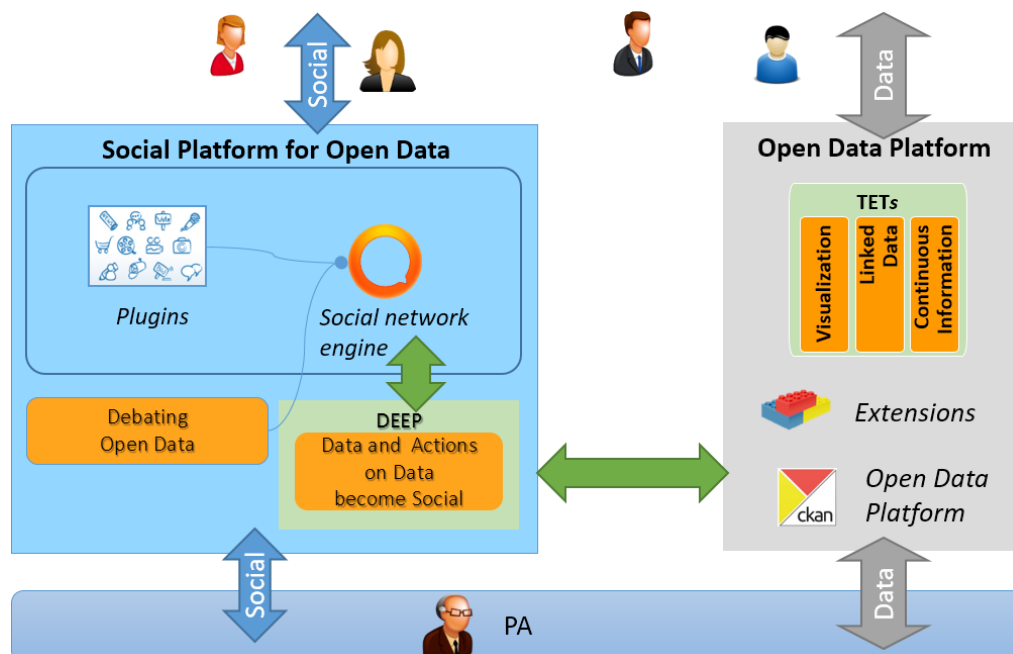


Figure 1: A general overview of ROUTE-TO-PA project: SPOD on the left side and TET on the right side.

SPOD (shown on the left side) is an infrastructure that includes a social platform component, and an open source, extensible, modular and pluggable architecture DataEt-Ecosystem Provider (DEEP), providing datalets, which are web-component visualisations of datasets content, that enable the sharing and collaboration around data. Among the services offered to citizens by SPOD we provide an easy navigation of different open data

providers with the same interface, allowing them to browse the archive interactively, a personal space (MySpace) for reflections and for taking notes on visualizations, URLs, and texts, a set of Public Rooms (in an Agora) that allow public discussions and (in the Beta) a Collaborative Space (available only in Beta version, end of year 2) where small groups, privately, can discuss and interpret collaboratively the open data and make sense out of them.

TET, on the right side, is a CKAN extension that aims at supplementing the CKAN functionalities by providing information about the quality of published data and review of datasets based on usage experience and linking of related datasets to improve understanding of underlying subject matter and issues. It also offers personalised presentation of datasets based on profiles and preferences of users, and extraction of implicit patterns in published data for greater insight, in addition generating examples of data violating specific rules, such as anti-corruption rules.

SPOD and TET are strictly intertwined as the datasets provided by TET with usage experiences will be employed by SPOD to present relevant datasets to enhance and improve the discussion and viceversa, social interactions will be used by TET to improve the personalisation of data. These advanced features are only prototyped in Alpha version and will be developed more completely in the Beta version (end of year 2).

Our technology is meant to address some fundamental issues about the citizens and open data. Those issues have been identified in the activities of WP2, with the final result within the D2.4, with a list of Use Cases (UC), that have been used as functional requirements for the design of SPOD (D4.1) and TET (D4.2).

It must be noticed that, of course, only some of these features are implemented in version alpha of SPOD and TET, but the underlying architecture and their full integration is crucial to ensure that citizens' needs are addressed by the ROUTE-TO-PA platform seen as an ensemble of connected and seamlessly integrated pieces of software.

So, pre-requisite is the **integration** of the two platforms since the very beginning of the design. We have provided a joint authentication mechanism (based on open protocol such OpenID) (see D4.1, sec. 2.2.1, req.3 in the table) that allows a citizen to be a user recognized in the same way both on TET and on SPOD. It allows also the exchange of information between SPOD and TET (as described in D4.1, sec. 5.3) that support the operations described below.

The first issue to be addressed for citizens is **searchability**, i.e., the capability for a citizen who has a question, or a problem, of looking for suitable datasets. The issue is well described in the D2.4, in Use Cases UC11: *Query data* where it is required that TET allows to query a dataset to obtain a correct or useful sets of data for analysis, and UC 18: *Receive notification for New Datasets* where it is described how the user receives notifications of new datasets, i.e. from SPOD friends, or a specific topic s/he is participating. But, it is also important to provide to citizens the ability to explore all the datasets that are available. This is supported by the requirements of SPOD that offer the exploration of datasets with an interactive navigation (see D4.1, sec. 2.2.1, req.16, 40 and 45 in the table) of the associated dataset providers (see D4.1, sec. 2.2.1, req. 44 in the table), i.e. those repositories of datasets that are of some interests to the community and that offer (in general) relevant datasets. The administrator can set up these associated dataset providers in order that citizens are easily provided with the exploration and search among *relevant* datasets, including, for example, the open data portals of the town, province, country, the national statistics office, the european data portals, some open data portals that offer datasets of specific interest to the community (for example, suggested by the hot topics for citizens, like demographic information, or a specific industry or commerce area, etc.). Moreover, both SPOD and TET are required to explore how the system can automatically generate a list of recommended datasets based on user's profile UC19: *Request recommendations for Data Sets*, where citizen's profile (as used in TET)

combined with the social information of his/her network of friends in SPOD are used to provide such as suggestion.

Finally, of course, both SPOD and TET offer the standard way of searching among datasets by term, keyword, etc. in order to provide basic functionalities to find appropriate datasets.

The second main issue is **availability, suitability and usability** of datasets. Our design addresses this issue by providing tools to the citizen to identify, select, filter datasets and preview the results in order to guide him/her through the choices and the selection. For example, in SPOD, when selecting the type of charts to be used, a preview immediately shows which kind is not suitable for the data. TET offers similar support to users that are also offered the possibility to filter out and make simple elaborations of the dataset (standard operations to count the number of occurrences or use the max/min of a column, for example).

It is important that ROUTE-TO-PA has a crucial focus on supporting the citizen in dealing with (possibly complex) datasets. It is well recognized in D2.4 in UC20: *View Chart* and UC21: *View Infographic*, where in the “Special requirements” it is explicitly mentioned simplicity and user friendliness. Also, in D4.1, the non-functional requirement of Usability is taken into account (D4.1 sec. 2.1.1.3) and used, for motivation, for several functional requirements in sec. 2.2). In the design of our platform we plan to devise a mechanism such that, when a user selects a dataset, some suggestions are provided about the best way of visualizing it. First, being SPOD a social environment, a source of suggestions will be the choices previously made by the other citizens on the platform. So, a suggestion like “Most popular charts for this dataset...” will be shown according to the statistics. Then, some heuristics will be used on the dataset to suggest alternative (equivalent) visualization charts, as well as suggesting new ones.

LinkDaVis is an example of tool to support non-technical users in finding suitable visualisations for the specified dataset. In particular, it automatically categories and binds the data to the visualisation parameters. In addition, as example of heuristic to automatically extract and determine the data type categories, the on-line system Many Eyes created by IBM (but since June 12, 2015 not available anymore), analysis the uploaded dataset categorizing the data type content and recommending appropriate visualisations.

The issue is also dealt within the design of SPOD, mainly, because of the non-functional requirement of Interoperability (D4.1, sec. 2.1.1.2) where the ability of SPOD to read dataset from a wide range of providers is required. As a result, the SPOD system is built so that the datasets can be loaded not only from TET, but also , sec. 4.3.6) from any CKAN server (that has activated the features of remote APIs), from OpenDataSoft, from UltraClarity, and in future with other protocols like Oasis ODP.

The third issue, is the **quality** of datasets. This is a crucial issue dealt with, in our analysis, by UC2: *View provenance and Metadata related to a Dataset* and UC3: *Check and rate Dataset quality* in D2.4. TET (based on some mechanisms in CKAN, but significantly improving on that) will be offering several quality parameters based on well-known and widely recognized Open Data guidelines, including metadata, provenance, context, last update, frequency of update, supplier/publisher’s name, contact name, etc. This will enable the SPOD to offer measures to citizens about the quality of the data they are referring to, but also to feed back to TET, citizens’ opinions on the quality of the data and how useful it was during the discussion (as described, for example, in D4.1 sec. 5.3.1, DataFlowSpod3 in Table 8).

3 THE SOCIETAL LEVEL OF ACTIVITY

Work package 3.1 resulted in the development of the Societal Activity model of Open Data use. The purpose of the model is to enhance our understanding of the user requirements of open data in a societal context. The model helps to find the best fit between; on the one hand, the impetus for governmental organizations to provide open data, to increase accountability and transparency, and the specific needs of citizen-users in particular domains. This involved representing what is at stake for open data provider organizations, the processes by which this data is (or is not) made available, the motivations and organizational representations of this activity, and, on the other hand, understanding the needs and motivations of citizen-users. The model followed a top down approach. It identified three democratic traditions: monitorial, deliberative and participatory. In each tradition both citizens and public administrators have different roles and accompanying needs and motivations. At the organizational level, the varieties of transparency were examined and the concepts of participation and collaboration as part of an open government were explained. Both the democratic traditions and organizational components were taken into account as being part of the context in which the activity, the usage of open data, takes place. The usage of open data by citizen users and public administration-users were considered as two separate activity systems, interacting together in a network. This resulted in the Societal Activity model of Open data use.

The model was developed based on the literature and tested in five pilot sites, using interviews, analysis of official documents (where available), and workshops or focus groups where open-data providers and users met and discussed. For each pilot site the different components were identified before the development of TET and SPOD. In addition, it is important to be aware of possible disturbances in the model. Engeström (2001, p. 135; 2000, p. 960) stresses that (the overcoming of) tensions in the activity model are key to process of learning, and function as the driving force of change and development. Tensions or challenges across the pilot sites were identified. Most prominent are the challenges within the community of the citizen users and within the community of the public administrator. Additionally, across the pilot sites challenges regarding open data can be identified, which relate to access, quality and the lack of data. Furthermore, ownership of the tools, moderation of the tools and trolling are mentioned. Lastly, there might also be tensions between the two interacting systems. For instance citizen users might be interested in data that conflict with privacy issues or that PAs fear will have a backlash. These challenges are important not only for the development of TET and SPOD but also for the evaluation because they can be observed over time and possible changes can be evaluated.

Work Package 2 (WP 2) and Work package 3 (WP 3), provide the theoretical and empirical foundations for setting-up, understanding and evaluating participation in technologically supported activities around open data. WP 2 provided input for user requirements. Work package 4, the design of TET and SPOD, provides technological requirements. The societal activity model (D3.1), one of the components of WP 3, adds two more levels: the societal level and the organizational (community) level. The latter will be further developed in WP3.2, focused on the interactive activities in Open Data use. This model will elaborate and visualize the participation within the community, including roles and clusters identification.

For the evaluation of WP5 we have identified four levels that are all interacting together (see also Chapter 2). For each of these levels, success criteria can be developed. The components of the Societal activity model are in line with the different components of the evaluation:

1. Technological criteria (Tool, WP4)

2. User criteria (Subject, WP2)
3. Organizational criteria (Community, Roles and Rules, WP 3.2)
4. Societal criteria (Object – Outcome, WP 3.1)

For the evaluation of TET and SPOD this means that the model can be used to study the dynamics in the societal activity model; between the citizen-user and pa-user. It can assess the process of open government efforts at the different pilot sites based on the components of the societal activity model, thereby identifying possible shortcomings, challenges and transformations over time. In addition, the outcome can be assessed. The outcome of the societal activity model is different for the three democracies.

4 THE COMMUNITY LEVEL

4.1 COMMUNITY BUILDING

The success of the ROUTE-TO-PA project depends on having stimulated the emergence of active and perennial online communities of practice, mediated by (SPOD-TET) tools that both facilitate activity and whose innovativeness attracts new users. A significant amount of effort will be expended in the project in order to accompany and favour the emergence of such communities.

Over the last 20 years, a research literature has been developed on the study of the emergence, functioning and evolution of online communities of practice (e.g. Wenger, 1998; Wenger et al., 2002), and on strategies for stimulating their growth (e.g. Kraut & Resnick, 2011). Maloney-Krichmar and Preece (2005, p. 203) define an online community as a group of people with a *common interest* or a *shared purpose* whose interactions are governed by *policies* in the form of tacit assumptions, rituals, protocols, rules, and laws supported by technology. There are different types of communities e.g. discussion communities but also task or goal oriented communities and combinations (Stanoevska-Slabeva & Schmidt, 2001).

Typical online communities go through a series of developmental stages, each of which requires adapted community building strategies (Wenger et al., 2002):

Community developmental stages

Stage	Characteristics
Potential	Movement of networked individuals towards becoming more connected and a more important part of the organisation.
Coalescing	Members interact and build connections. At this stage they form a community.
Maturing	Growth in participants and depth of knowledge shared. Moments of high and low activity.
Stewardship	Communities take active management of the knowledge and the practice they have and members are aware of how these are being developed.

Transformation	Activities and shared practices change. At this stage, communities can fade away, merge with another community, or die.
----------------	---

The following are some of the best practices for online community building that can be extracted from the literature (e.g. Kraut & Resnick, 2011):

Best practices ("BP") for online community building

<p>1. Community purpose. Define and articulate the purpose of the community; identify members' needs, through online surveys; use chat rooms to run focus groups. Develop a clear narrative, identify core values of the community; communicate the narrative through the layout, activity tools, and content of the website tools.</p> <p>2. Collaboration spaces. Build flexible, extensible gathering places that can evolve with the needs of the participants. Start small and allow development over time.</p> <p>3. Profiles. Create meaningful and evolving member profiles, that will allow participants build trust and social identity, foster relationships, and give the community a sense of history and context. Track and display participants' usage history.</p> <p>4. Roles. Design for a range of roles, to be managed by a site moderator: Visitors need to be welcomed; Novices need instruction; Regulars need rewards; Leaders need to be empowered; Elders need to be appreciated and honoured</p> <p>5. Leadership. Encourage the emergence of strong leaders, usually on the basis of "recommendations" from other participants, in order to gradually devolve responsibility to the community itself (the "stewardship" stage of development).</p> <p>6. Rules. Encourage appropriate (n)etiquette, or community rules and values, especially for conflict resolution. FAQs, a privacy policy and guidelines for community leaders (or "admins") are important).</p> <p>7. Events. Promote regular events, including those organised by participants, including online group meetings, presentations and conferences, competitions, a calendar system, and chat rooms.</p> <p>8. Rituals. Integrate the rituals of community life, such as personal transitions towards roles, or particular milestones and success stories.</p> <p>9. Subgroups. Facilitate member-run subgroups, with appropriate technologies for managing them, in order to enable the community to grow.</p> <p>10. Content and tools. Develop attractive content and tools: many online communities are successful because they have a rich inventory of content, and/or tools for working on it, that attract new members</p>	
---	--

In accordance with theses best practices for community building, the ROUTE-TO-PA approach is as follows:

Best practices	ROUTE-TO-PA approach
1. Community purpose.	<ul style="list-style-type: none"> Members needs and communities' purposes identified in WP2 and WP3 studies; these will be further refined throughout the project by qualitative analysis (WP2, WP5) of emerging and evolving needs clear narratives of communities' purposes, identities, will be developed, in collaboration with initially identified core user groups (focus groups), in relation to SPOD-TET tool design
2. Collaboration spaces.	<ul style="list-style-type: none"> the SPOD tools allow creation of rooms focussed on specific topics for specific (e.g. P.A.) groups tools will be developed for identifying emerging spaces for deliberation and close collaboration (Social Network Analysis – WP2) and displaying their form to users, as reflection/awareness tools

3. Profiles.	<ul style="list-style-type: none"> each user of SPOD-TET tools will have a personal profile page, inspectable by all validated community participants
4. Roles.	<ul style="list-style-type: none"> site moderators will be appointed (from PAs, from research groups) who will, in negotiation with participants, identify sets of roles (such as “power user”, “admin”, “expert on xyz”, ...) and negotiate their ascription to participants moderators and users with related roles (e.g. admin) will define functions to be fulfilled – e.g. welcoming new users, providing help, rewarding high users — and implement them
5. Leadership.	<ul style="list-style-type: none"> (as above, BP4) moderators will gradually give over responsibility for management/moderation to participant-leaders, on specific themes, enabling devolution of responsibilities to ensure continuation of community after the project duration
6. Rules.	<ul style="list-style-type: none"> the project will define guidelines for carrying out moderator roles an initial set of community rules will be defined, inspired by existing online communities of practice (e.g. PYTHON, Wikipedia), and negotiated with participants a part of the moderator role is to ensure application of the community rules moderators will apply conflict resolution strategies where necessary (e.g. proposing that protagonists summarise the main evidence for their views; proposing a vote)
7. Events.	<ul style="list-style-type: none"> workshops, face to face, will be organised twice a year by all partners, in a “hackathon” style, on specific open data themes of interest to the particular communities (see WP2, scenarios)
8. Rituals.	<ul style="list-style-type: none"> WP2 and WP5 work on evaluating the form and nature of activity, collaboration, using the SPOD-TET tools will enable identification of achievement of key milestones of communities (e.g. attaining a specific number of members, number of open data visualised), which will be fed back to the communities themselves key participants will be encouraged to organise specific meetings
9. Subgroups.	<ul style="list-style-type: none"> the SPOD tools allow separate rooms for specific subgroups moderators will encourage subgroups to define their mission statements and rules for membership
10. Content and tools.	<ul style="list-style-type: none"> “Content” of the online communities concerns both the open data provided and the visualisations of it. Ensuring content that attracts new participants therefore depends on: (1) efficient negotiation between user-participants and PA-participants, with the latter providing new content to meet community needs. A negotiation process workflow will be implemented, inspired by the PYTHON online community, that involves proposing new content, identifying a “champion” to argue for the necessity of that content, other supporting arguments, and a means for clarifying and recording outcomes usability work in the project will ensure that tools are beyond the state of the art; this will be reflected in the communities’ and the ROUTE-TO-PA tools’ narratives.

As can be seen from the above, **the moderator role is crucial to success** in community building. During the project itself, moderators will be both PA and research institute representatives, who will work in the manner described above to devolve responsibility and empower community participants to take over these roles.

Dissemination activities will all contribute to the engagement of new community participants.

The literature highlights that in virtual communities there is usually 1% of the participants highly active whereas the large majority only follows debates and does not actively contribute by adding information. These “lurkers” read discussions in a newsgroup but rarely or never participates actively (Nonnecke & Preece, 1999). In addition, there seems to be an intermediate group – approximately 5% – that poses their own new questions occasionally but hardly participate in other debates. This group is important for broadening up the scope of the debate (Meijer et al., 2012). For the management of the online community the 1% and the 5% are crucial for

the content generation but the large majority – 94% - is crucial for the broader dissemination of the content in society.

4.2 FEATURES OF COLLABORATION

Within the community at some stage of development, some participants are required to collaborate in order to work with open data and jointly discuss issues or solve some actual problem. For example, in the Den Haag case, employers and public administrators, the first representing the (highly diverse) local community of employers, the second representing local government, have decided to use the tools to find solutions for local problems of employment and regulations about employment. Open Data are used to depict the current situation in various parts of the municipality.

Getting adults to collaborate will not just happen, even when they are presented with a great collaborative tool. Not only do people have different conceptions and experiences in the collaborative realm, also may it be the case that people's interests, roles, autonomy, and most importantly, their individual objectives differ. As a consequence, there will be insufficient trust, unclear expectations, ignorance of steps to follow, each of these issues being sufficient reason for collaboration to fail (Allwood, Traum & Jokinen, 2003).

Also in contrast to what many people may think, social media such as Facebook or Google+ do not attract much extended dialogue (de Villiers & Pretorius, 2013), they are mainly used for socialising (Mazman & Usluel, 2010) and have little educational use (Hew, 2011). The contributions of most members are simple announcements, and some immediate comments on these announcements, but rarely more than a coherent exchange of a few contributions. Moreover, such contributions are rarely argumentative, or even explanatory. And finally, in spite of what many people believe, it is very hard to regulate emotions in social media, especially asynchronous media (Andriessen, Pardijs & Baker, 2013).

For these reasons it is imperative to consider a number of theoretically and empirically supported foundations of collaborative interactions using social media for knowledge construction.

4.2.1 COLLABORATION

For describing what people actually do rather than what they are thinking or thinking they do we need to analyse their activity when using the tool. At this level we expect most of the learning to occur. However, simply presenting the medium to users does not result in the desired activity, and not all activity is collaborative activity, and not all collaborative activity leads to the required goal. Therefore, some training (and moderation) in these aspects is strongly advised. Below we describe an approach for achieving a basic level of collaboration in social media for our particular purposes.

Collaboration and learning evolve and should evolve along the following six dimensions of activity (Andriessen & Baker, in preparation):

- Dedication to the task (cognitive consideration): or the degree to which the user is serious about the objectives, shares most of the objectives with the others, and can relate in some way to these objectives and the group achieving them;
- Understanding the process: evaluating and assessing what is going on and what that means in terms of collaboration, progress towards the goal, and required action; knowing how to deal with discussion in an asynchronous medium;
- Need for support: Knowing when to ask for help, recognition that one is stuck, or the obstacle is difficult to surmount, and knowing whom to ask for help, including knowing to explain what is going wrong;

- Linking the activity to what the user already knows, but also which maybe is the most important aspect of learning: creating an emotional bond between learner and learning object and/or activity; especially unlikely connections are good for generating creative solutions;
- Abstraction: Knowing to switch from concrete to abstract and back is a good indicator of metacognitive understanding;
- Ethical consideration: understanding others need support and acknowledgment for their actions and contributions, and providing that support; this is a crucial aspect for the development of trust in a collaborative team.

Some of these dimensions are especially important and difficult in social media. In a social medium it is very important to manage one's silence. While in real life any reaction and interpretation of one's activity or utterance can be read off the other's face immediately, in a social medium, especially an asynchronous one, the most immediate reaction always is silence, nothing happens. That is why we need to give users an idea of **knowing when to act**: the need for the group there always is a timely reaction to new contributions.

The main outcome of all this activity is called **trust**, the team can rely on each other and knows that it can. Another outcome is **learning**, which needs to be defined in terms of the goal of the activity, or study one is engaged in. A third goal, in this project, is a contribution to increased **transparency**, at the community level, the communities are those which our participants are a part of. The definition of this may also be related to the type of activity. For transparency to be a main outcome of our work, it needs to be particularly stressed as a goal of user activity every time, which means we must be explicit about what we expect our users to do in this respect.

4.2.2 PHASES DURING THE WORKING PROCESS

To create consensus on the important features of working with the tool, we will try and describe general phases in this process, and try to describe the differences from a main working model for each of the cases.

- **Issue finding**: We assume each of the scenarios revolves about a relevant main issue or question. This is a very important phase, and not all issues lend itself for discussion or discussion with open data. An issue needs to be relevant for the participants, meaning they will have sufficient knowledge to deal understand the issue and be able at least to bring forward some relevant ideas. Furthermore, the outcome needs to have some interest for them, ideally speaking, it has to be an outcome that will be used in further discussion on policy making. In this sense, we can say we are involved in a dialogical process (Paavola, Lipponen & Hakkarainen, 2004). Another requirement would be that the users (and the further elaboration of the issue) have an interest in data, and not (solely) in opinions. Therefore, the participants are interested in relevant open data. We may assume that not all know where to find the data, how to use them, or what data would be relevant in the first place. Therefore, the phase of finding open data is never trivial, whatever the level of sophistication of the tool. Working with open data is part of the users' learning curves.
- The second phase would involve **finding open data**. This could be an individual effort, supported by computer tools, such as TET. The user tries to look for data that will illustrate a particular point he wishes to make concerning the issue. But it could also be the case that the data are the issue for discussion, because they represent some development or result. The process of finding open data needs more attention in this project. We assume that the stage of finding open data, includes

reflecting on the relevance of what is found, and storing the open data in some space, such as the private room on SPOD.

- The third phase is **visualising open data**. This requires a space in which some visualisation can be selected, for one or two variables. Many open data are not immediately clear about the variables and values they contain. In principle, this is the role for TET.
- The fourth phase would be **sharing the data**. Here collaboration becomes manifest, we therefore need a space where the data can be shared, with a comment or question. The space should be labelled by its initiator, and the central question should be explicit and clear. This space should be large enough to allow much room for additional comments, questions, arguments, conclusions. Because we cannot visualise more than one or two variables at the same time, there should be space for additional visualisations, within the same space, or on a next page (e.g. wall of the room). People can start more than one room, but inactive shared rooms should be closed down or removed to background.
- The fifth phase, which is most crucial, involves **discussing about the data**, within the constraints of the issue proposed by the initiator. The initiator can choose the discussion to be open, or closed, in which case participants should be invited. Discussion is about the point the data make concerning the issue. They should lead to a conclusion about the issue, whereby the conclusion needs to be shared by all. It can also be that the conclusion is not the final answer to the issue, but requires further elaboration on the next page.

The fifth phase is the most complicated one, in terms of collaborative activities, and also here we expect most differences, for example in discussion styles (see below). In simple scenarios, this could mean a public administrator answering an informative question by a citizen, by pointing at a feature of some open dataset. In more complex scenarios, collaborative exchanges greatly varying in length may ensue. One of the outcomes of the project may be a guide on what types of discussions work well with the tools, and how to engage in them, and what types of support (by a moderator or by the tool) would be required. Meeting community-level goals also involves users being aware of the community needs, and reflect on their collaboration in that respect (see 3.2.4).

4.2.3 DISCUSSION STYLES

In addition to the type of collaboration being sensitive to the phase of working, it also has different structure, norms and values according to the style of the discussion. Discussion styles are characterised by different characteristic types of contributions, and also different pacing, frequency and level (concrete and specific, general) of contributions. Obviously, they serve different (sub)goals during a discussion. Awareness of them would be part of what we could call interactive consideration, but the extent to which this is truly necessary, or an evolved outcome of goal awareness, remains a topic for further research.

A most promising style is called **Arguing to Learn**, or the use of argumentative discourse, that is giving reasons for or against some point of view, and backing up these reasons with information and data, not for the purpose of winning a debate, but with the clear goal of exploring and deepening the domain of the issue (Andriessen & Baker, 2014).

Another discussion style is Co-Creation, where the focus is on sharing and interpreting new information in the light of a joint question, the solution of which is the trigger for participation.

A somewhat less challenging style can be called **Information Sharing**, when there is no joint issue to solve, but participants have an interest in informing each other, perhaps based on differences in expertise. A more

interesting form is to achieve **Shared Understanding**, which needs to be checked and affirmed in some sort of collaborative dialogue.

The style of discussion will constrain the evolution across the dimensions mentioned above, as well as the types of outcomes. Hence, evaluations of the quality of collaboration and outcomes will be different, or will need to have a different focus, for example:

- For arguing to learn, it is less important that users produce correct arguments, but more important is it that they arrive at a deeper understanding of the domain at the end. This could be reflected in a strong conclusion, shared by all participants. During the process, we think constructive contributions are most important, that is, arguments, proposals, new information, and explanation, and at the social level, ethical consideration (see above). This is especially reflected in the ways counterarguments are formulated, and how others react to them. Hence, one does not say: I totally disagree, because.... but one says: That is very interesting, but maybe
- For co-creation, the focus is on generating new ideas and solutions, supported by relevant information and interpretations of that information. The appropriateness of a solution partly depends on the extent to which all relevant information has been considered. Here also, the composition of the group is important, as the conclusions drawn should be shared by members of the community, and, most importantly, all sides of the issue may require different types of expertise. For co-creation, a new contribution must be read, understood and shared by all participants, and all should monitor this will be the case. Obviously, such an approach requires high degrees of cognitive and ethical consideration.

What is important here is for every discussion to be clear about what kinds of contributions are expected and desirable, and to be aware if the rules of the game: what should and should not be said and done. Maybe, when groups are not experienced or the rules are somewhat unfamiliar, a moderator can help to bring forward relevant activities. Especially when people are not very active, some support will be useful.

4.2.4 USE OF OPEN DATA DURING DISCUSSIONS

Figure X: Screenshot from Google Gapminder



The societal model of Open Data identified a some potential tensions across the pilot sites regarding open data (D3.1, see also section 4.1 below). Relating Open Data to constructive discussion or information gathering is not an obvious action. Not all issues are suitable for this type of problem solving. People are not used to back up their ideas with data, and data are also often misused for various reasons. And even if the discussion requires them and people are understanding how to use them, Open Data may be hard to find and they

come in different formats. How far can we go to accommodate users? How to check if data are well understood? Above an example of an excellent presentation of Open Data, by google gapminder. The data are

even dynamic, and show the growth of the population and their income over time in one graph. Obviously, such data can be useful in dedicated discussions, but most data are of a different kind, unsuitable for fancy presentation. This also points to the limitations of depicting only one variable at the time. As the example shows, for some ordinal variables scaling is possible, but most variables are nominal. Resolution of such technical issues are still being investigated.

Research needs to be concerned with making users aware of the kinds of issues that can be researched with our tools and how to deal with the limitations of open data. For example, it may be the best solution to appoint a local open data officer for finding and visualising relevant open data, when asked by a participant in a recognised forum. Local administration could be aware of the problems by developing a policy for generating open data in a particular, annotated, format. Another solution could be moderation, where the moderator also supports finding, presenting, and interpreting open data, on demand for the users.

The most important feature of using and presenting open data with respect to collaboration is that the presentation and design foster users working in *design mode* (Bereiter & Scardamalia, 2003). Design mode can be contrasted with belief mode: in the first, the interest is in creating conceptual artefacts, meaning that users discuss concepts, ideas, related to conceptual goals, such as increasing transparency, and produce ideas that are really beneficial for others in the community, serving more general levels than the immediate, concrete needs. Questions that are asked about open data could be: *what are these data good for? What does it allow and not allow to assert? How does this relate to our issue? Will the community benefit? What other data could be useful?* In belief mode, users discuss together and produce solutions for the sake of producing something, not something necessarily useful or innovative. The main interest of belief mode is what people believe or ought to believe: (dis)agree, argue, express and resolve doubts. To support working in *design mode*, some features of the tool need to be further studied: support working with ideas towards further understanding, allow further elaboration and divergence of ideas for the benefit of a more general approach, to support users engage into regular summative reflections, in other words, to immerse the users in advancement of knowledge. To conclude, at the community level, effective working with data is not as straightforward as many may think. The power of collaboration is the possibility to jointly arrive at an appropriate solution by negotiation of interpretations. Such negotiation processes are crucial for learning and understanding as well (Baker, 1998), but not just any negotiation is a useful one, we need to have the higher interest of transparency in mind. Collaboration may be powerful when it requires a motivation at the community level to become useful as an activity.

4.2.5 MODERATION

As we have seen above, moderation is crucial to both structuring and realizing a good conversation and preventing forms of abuse, undermining behaviour and trolling. Lampe & Resnick (2004)¹ highlight that various methods can be used to limit the disruption that anti-social behaviour can cause and to help the participants to cope with information overload. While the traditional perspective highlight that there is one or several moderators that are appointed by the 'owner' of the website, currently more distributed approaches to moderation have been deployed. These approaches stress that people can become moderators when they earn certain points and their rights as moderators can increase. In smaller communities, however, moderators will need to assume a more active role to get the conversation going, attract new participants and stimulate debate and collaborative work. This is key to bringing the community to life.

¹ Lampe, C., & Resnick, P. (2004). Slash(dot) and burn: Distributed moderation in a large online conversation space. Proceedings of ACM Computer Human Interaction Conference 2004, Vienna.

In the collaborative group, operating asynchronously on a social platform, moderation of idea generation, collaboration and reflection is not an obvious task. It cannot be simply based on results from instruction, or synchronous forums (Asterhan & Schwarz, 2010). Requirements could even be contradictory: moderators are expected to save the group out of a (conceptual or collaborative) predicament, while this same action precludes important learning gains for the group itself, when it has solved the problem by itself. This means that each group needs to negotiate the type of moderation that is needed. A moderator who solves all problems for the group is not a good moderator, and a poor teacher as well.

4.3 HOW ROUTE-TO-PA FACILITATES COLLABORATION

The list below concerns our current line of thinking on these issues. Actual implementation depends on the expertise and objectives at each user site, and on the final tool affordances of the alpha-version.

Best practices	ROUTE-TO-PA approach
1. Collaboration	<ul style="list-style-type: none"> • Organise preparatory workshops for potential users focusing on awareness on good collaboration • Provide moderator – based feedback during working with the tools • Design small activities creating awareness of collaboration • Co-design meaningful assignments and objectives, and be clear about the desired level of collaboration • Allow collaboration time to evolve, preferably in rooms with a limited number of dedicated users
2. Phases	<ul style="list-style-type: none"> • Allow interfaces to look different or options to be available depending on the task goal and phase • Work with different scenarios and develop a GUIDE for different phases in scenarios
3. Discussion styles	<ul style="list-style-type: none"> • Design templates for guiding users to design the type of contribution they make in terms of a desired discussion schema • Allow overviews of contribution-types in meta-views • Stress information sharing by asking for frequent summarising activity
4. Open Data	<ul style="list-style-type: none"> • Users may not interpret open data appropriately, this needs to be monitored in the light of community goals and learning processes • Developing modules for using SPOD as a learning tool for interpreting open data are being considered • Carefully calibrate design features, instructions, tool feedback, and human support
5. Moderation	<ul style="list-style-type: none"> • site moderators will be appointed (from PAs, from research groups) who will, in negotiation with participants, identify sets of roles (such as “power user”, “admin”, “expert on xyz”, ...) and negotiate their ascription to participants • moderators and users with related roles (e.g. admin) will define functions to be fulfilled – e.g. welcoming new users, providing help, rewarding high users — and implement them • moderators support the group, but do not lead the collaborative work

5 ACTIVITIES AT THE USER SITES M6-M12

The current state of the work on the cases does not allow complete descriptions of the preparation phase. For some user sites, this will only start when the first prototype of the tool will be available. Nevertheless, users and researchers have been active during this period with many things.

5.1.1 PRATO

The pilot activity in Prato is currently focused on the usage of the SPOD/TET platform to stimulate discussions on the city budget issues by allowing citizens to access budget open data and other information available on the municipality open data platform.

In the M6-M12 period, activities were dedicated to the analysis of the local technological and open data context, in order to verify how to make budget data accessible through the SPOD/TET platform, and to the initial planning of the pilot operational context, in co-ordination with the institutional actions carried out by the administration for the approval of the City budget.

At the end of 2015, the Prato City Council has approved the harmonized estimated budget for 2015-2017. This is the key instrument for the planning and control of the city and the document reports the three-year forecast approved in overall balance between total income and expenses.

The municipality worked hard to gain such result and to give the city a consolidated financial tool for the administrative government and this represents an unexpected success. On the other hand, from the project point of view the initial ideas on the pilot activities had to be a little re-thought, since at the beginning the focus was meant on the usage of the SPOD/TET platform to rise discussions quite on the forecast budget, to support the administrative decisional process, while now the pilot action needs to be shifted to a post budget analysis discussion, with the aim of suggesting possible changes for the future.

Almost simultaneously with the start of the Route-to-PA project, the City of Prato started the publication of available open datasets by activating its own CKAN platform. Such platform was initially provided by the Open Data Network (ODN) project, co-ordinated by the Province of Florence and including the Province of Prato and Pistoia and the Basin Authority (www.opendatanetwork.it). Presently, the City of Prato node of the OpenData Network includes 77 datasets, the majority of which (40) are related to population data (odn.comune.prato.it).

A primary issue for the project was therefore the verification of the connection of the ODN platform with the SPOD platform, in order to make the Prato datasets accessible for the pilot activities. Meanwhile, to speed up things and allow the testing of the SPOD platform by the Prato test group, some of the budget datasets of the Province of Prato, already available on the ODN platform, were copied in the SPOD. The technical group in the Prato municipality responsible for open data publication was involved in the analysis of the SPOD available datasets to verify compatibility and the result was that some minor changes in the data format were required on this side.

Unfortunately a major problem arose when the technical staff in Prato realised that the DataStore APIs (necessary to give access to third parties) were not available on the ODN CKAN present version (1.8.1), therefore the SPOD platform was not able to interrogate the ODN CKAN. To solve this issue, it was decided to install an updated CKAN version including the DataStore APIs, namely the SPOD one, directly at the Prato Municipality premises and to transfer all the available datasets, including the forecast budget data 2015-2017, onto this new platform. In this way, a project dedicated platform will be available for the pilot activities and it will be easier to manage it without depending on the Province of Florence, responsible for the ODN platform.

The final installation and configuration of the new CKAN platform at the Prato Municipality is expected by the end of January and some extended pilot actions should start in February.

5.1.1.1 INITIAL TEST ON THE SPOD PLATFORM

Following the first release of the SPOD platform at the beginning of October, an initial test group was appointed to verify the platform functionalities and give feedback to the developing team. A first meeting of the group was carried out at the end of September and several comments were provided on the available platform versions, both by email and through the available feedback form.

Since the city budget data was not yet available, the test was based on the data of the Province of Prato already recorded in SPOD platform and was mainly directed to check the building of the appropriate datalet for data visualisation and their inclusion in the discussion on SPOD (Fig.), even if some links to the forecast budget data were included in test discussion (Figure 4):



Fig. 4 Example of datalet created with the budget data of the Province of Prato in the SPOD platform.

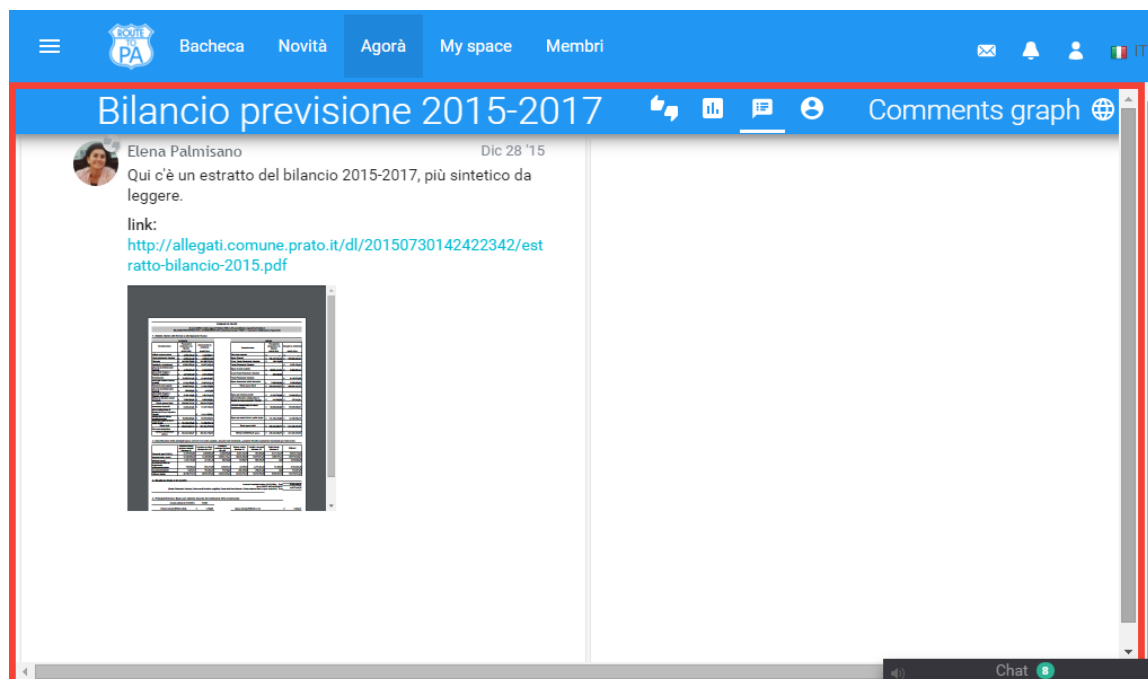


Fig. 5 Link to the forecast budget data of the Municipality of Prato added in a discussion on SPOD..

The members of the test group will also be involved in the extended pilot actions and their experience will be a valuable asset in the pilot case definition.

According to the schedule of the CKAN installation and configuration (end of January) and the final release of the SPOD platform, the first launch of the pilot activity is planned for the 11th of February, in occasion of a meeting with a restricted number of city stakeholders for the presentation of the SPOD/TET platform.

5.1.2 GRONINGEN

At the user and technology level the main aim for Groningen is the provide input for the design of the tools for open data use and to test the tools. Important for Groningen is that the tools facilitate discussion and collaborating regarding population decline issues. In order to prepare our participants in Groningen for using the tools we conducted a document analysis, interviews and organized two meetings of the "Population Decline project team in months 6-12 of the project.

The population decline project team consists of researchers, pa's and users and was formed in the fall of 2015. This project team will engage in the further development of TET and SPOD based on a concrete scenario regarding population decline. In addition the project team will develop success criteria for evaluation and will provide feedback for the design of the tools, resulting in better technology. Moreover this is the start of community building.

The group has 10 participants, consisting of the researchers as moderators, pa users and citizen users. Some participants also participated in the workshop regarding the user requirement but others were new. Next to pa's we invited people who are directly involved in citizens initiatives regarding population decline.

In the fall of 2015 two meetings were held: the first meeting focused on issue finding and the second on finding open data in relation to the issues identified. In order to prepare for the first meeting a document analysis was conducted consisting of government strategic plans regarding population. Reports of independent organizations and websites were studied as well.

The goal of the first meeting was to align possible interesting issues for pa's with the interests of citizens groups that can be discussed on the platform. The issues needed to be relevant for both groups and rich enough so that it will provide opportunity for discussion among the groups. In addition, during this meeting it was discussed how TET and SPOD can be tested and by whom. This first two hour meeting consisted of five participants (pa's and researchers) and resulted in a focus on "housing, labour and health care policy issues" within the scenario population decline. The researchers were in charge of the program of the meeting and we suggested topics to discuss based on an earlier performed document analysis. We did not record the meeting because we thought it might influence the dynamics. Extensive notes were made.

Following interviews by phone were conducted with three possible citizen users and a conference in Groningen focused on population decline in Groningen was attended. The interviewees were selected because they are involved in concrete projects identified during the first meeting: housing, labor and health care issues in relation to population decline. During the interviews the project was briefly introduced, following users were asked about to describe the citizens initiatives they are involved in, their relationship with pa's and whether or not they would be interested to participate in the ROUTE-TO-PA project. The notes of these interviews served as input for the second meeting with the project group.

In addition a conference on population decline was attended. The aim was to get an idea of the context in which the users operate, what the main issues are regarding population decline and what their motives are. Extensive notes were taken during the conference.

The second meeting of the population decline project group was held with 10 participants consisting of researchers, pa's and the interviewed users. Together two concrete scenarios were developed based on the experiences and challenges the users experience daily in their projects. For each scenario a description was given, information needs and relevant datasets were identified. It was noted that these datasets are currently scattered among several organizations and hence of the first aims is to find these datasets and place them all on one portal, CKAN that allows for TET and SPOD to function. Currently the Province of Groningen does not yet use CKAN. The researchers were in charge of the program of the meeting and suggested topics based on the interviews and document analysis conducted. We evoked discussions between the members but did not intervene much, other than summarizing. We did not record the meeting because we thought it might influence the dynamics. Extensive notes were taken. In the next meetings of the project group in year two, we will focus on third, fourth and fifth phase: visualizing open data, sharing the data and discussing the data.

5.1.3 DUBLIN

As reported in D2.2, D2.3 and D3.1, the Dublinked use case fits into a deliberative democracy model, facilitating civic contribution to debate on policy issues, specifically to encourage community conversations and build community networks. The Policy Issue is building community awareness and capacity for engagement. The motivation for communities to engage is the chance to find out what is already available/happening in their neighbourhood and what is planned for their neighbourhood in the future with the ultimate outcome to 'make my area great/make my city great'. The community of users involved includes local authorities, elected representatives, communities of interest, including residents associations, community groups, interest groups, and individual citizens. The Role of the local authority is *facilitator* and *data provider*, and the other community players are *participants* in community conversations and network building. The use case is supported by the Dublinked initiative.

Through the Dublinked initiative, the Dublin city region opens up access to open government data from transport, planning, environment and spatial mapping that are produced through the delivery of local services.

The greatest challenge for any data portal is not the publication of data but encouraging the use of such data. To encourage reuse, Dublinked combines an open data portal with the promotion of an innovation network to enable new collaborations that solve urban challenges and create better services. Dublinked user engagement to date has included thematic ‘dubmeets’ ‘hackdays’ and ‘open data challenges’, which have mainly involved businesses, technology developers, researchers and city experts. This has fostered a data innovation ecosystem that demonstrates the added economic value of open data through the creation of new data driven businesses, products and services.

Dublinked is now moving into a new development stage, seeking to broaden its user base and explore the potential social value of open government data. One issue is how to better communicate the value of open data to a wider audience and proactively involve ordinary citizens. The Dublinked data portal provides data that is generally accessible only to the IT literate, not to the general population. Dublinked offers a variety of search tools that examine the contents of the files and not just the metadata. However, with over 300 datasets from various agencies in the datastore, the challenge remains in finding data relevant to the untrained user, and in finding appropriate linkages between datasets. The second challenge is how to foster new types of community engagement around city data and create value for general users who may not have the technology skills to interact with the data available in raw formats.

The key objective is the development of new transparency enhancing technology that:

- Enables citizens to easily find data that is useful and relevant to them so they can find out what is happening in their area
- Enables citizens to explore and visualize a combination of relevant data to gain new insights about their city and neighbourhood.
- Enables citizens to have an easy platform to make their voice heard and get information on areas of interest.

Subsequent to the collective intelligence and user needs workshop conducted in April 2015 and reported in D2.2 in July, a variety of activities have taken place to advance the goals of the Dublin case. Throughout July to December 2015, there was intensive engagement as follows:

- Engagement between Dublinked and the Galway team to translate identified user needs into user and system requirements for TET.
- Engagement with the team in Utrecht to feed into on deliberative democracy societal model report (Deliverable 3.1).
- Engagement with Salerno team to translate user needs into user and system requirements for SPOD
- Feedback on SPOD features and usability from core research team in Galway
- Engagement with the team in Warsaw in developing the SIM model
- Preparation of demographic profile of Dublin region

Additional work preparing the evaluation of SPOD and TET and integration the Dublinked platform was carried out between November 2015 and January 2016. This involved the development of a new Wordpress website to integrate TET and SPOD and planning how to build capacity for wider engagement of end-users around open data.

A User focused design was adopted for the new DubLinked platform and data were curated to satisfy the needs of the identified user personas:

- *Government – the goal for this user is the “Browse” feature of the TET to help in the decision making process. The available datasets are DCAT compliant and linked to data.gov.ie, data.localgov.ie etc.*

- *Researchers – this category of user are interested in the “Analyse” feature of the platform to easily explore, merge, analyse, visualise and tag with feedback from expert researchers.*
- *Developers - this category of users use the “Build” feature to build commercial and civic solutions and tags real-time data and available APIs.*
- *Citizen: This user category generally need to be informed and empower based on the information obtained from the dataset. They use the social platforms and forum to bring about change (SPOD).*

Further initiatives to help prepare users for active engagement with integrated SPOD and TET involved the development of a wider policy framework Nov15-Jan 16. Specifically, Smart Dublin has developed a wider policy framework to engage citizens and stakeholders in the co-creation of solutions to identified city region challenges, supported by and using open data. This work involved an initial challenge identification workshop with staff in the four Dublin Local Authorities, which highlighted challenges linked to:

- *Transport & Mobility –Promote intelligent, efficient and integrated transport systems and active travel*
- *Environment & Public Realm -Promote sustainability, environmental quality and enhancement of public realm*
- *Flooding & Emergency Response - –Better monitoring of pollution, emergency response and building climate resilience*
- *Energy and public lighting – Promote energy efficiency in public buildings, and for citizens and business users*
- *Better data management - Promote data innovation to improve policy, service delivery and decision making*
- *Better communication and citizen engagement – Use open data to promote participation, engagement and inclusion for all Dubliners.*

5.1.4 CNRS/ISSY-LES-MOULINEAUX

As reported in year 1 WP2 deliverables, and in D3.1, the CNRS-Issy-les-Moulineaux scenario focuses now on the activity of two central groups: young entrepreneurs in ICT domain who form a major part of the local economic base and public administrators from Paris Region who collaborate with Issy-les-Moulineaux in the global open data policy. By “young entrepreneurs” we mean people who are in the process of creating a company and people who have created their company less than two years ago. Public administrators are representatives of public affairs. They apply, supervise and coordinate the policies programme of governments at local and regional level. These two populations have different approaches towards the use of open data.

In accordance with policy commitments, PAs generate public access to policies and financial information with the aim of restoring public trust and thus facilitate dialogue between electors and local governments. They want to make government processes and decision open. Intrinsically, they seek to foster the development of an economic environment around the transformation of open data into new applications and services. In this way, they want to improve communication between local communities and private companies in order to find a converging model of development that boosts employment.

Entrepreneurs want to build sustainable enterprises to create a network of business and public contacts. More specifically, they are looking for public subsidies and they are interested not only in promoting job creation but also in making money.

The CNRS team has studied the activity systems of these populations during two focus groups. The first one was realized in May 2015 with young entrepreneurs from an incubator and the second one in July 2015 with PAs.

Focus group 1: Entrepreneurs point of view

Concerning the activity system of entrepreneurs (for more details see WP3 deliverable appendix B), the members of the focus group revealed their expectations relating to a *collaborative platform devoted to open data* (e.g. to be informed about public politics in their domain, share knowledge) and the limits of such a platform (not agreeing to discuss on information about design process of new products, financial and economical data, business strategy, and organisational frameworks).

Thus, this focus group brought to light *two principal tensions* in the present activity system. In a first instance, entrepreneurs seek to interact around updated data in a synchronous way in order to obtain/maintain a competitive advantage by co-designing products in a short time period; but the current tools do not enable them to do so. In a second instance, young entrepreneurs have limited financial and human resources. *De facto*, they need to collaborate with other companies to develop some applications, but they do not wish to disclose confidential information with societies that they could be in competition with in some lines of business.

After studying activity system of entrepreneurs during the first semester of the project, the CNRS Team in collaboration with Issy-les-Moulineaux focused on public administrators in July 2015, then evaluated the first version of the SPOD/TET in October-November 2015.

Focus group 2: PAs Point of view (July 2015)

Seven civil servants from the Paris Region participated in a second focus group (for more details see WP3 deliverable appendix B). They are representatives of geographic information systems, representatives of associative life (promoting creation and development of local associations) and representatives of communication services (communication actions towards the general public, media and partners cities). They belong to a special interest group, organised by the Issy PA, and with which the CNRS team collaborates, that deals with issues related to open data in the Paris Region.

In this context, the focus group highlighted that PAs are the principle data producers and thus has suggested that they would play a major role in the development of a collaborative platform devoted to open data (SPOD/TET). More precisely, it has showed that PAs collect data, classify it and choose what could be published or not. They generate public access to data on public finance, mobility and geographical information with the aim of restoring public trust and facilitating dialogue between electors and local governments.

Results of this session have also revealed Public Administrations want to stimulate the social economy and local development by promoting the use of open data as a resource for the creation of services and software tools.

Therefore, PAs seek to facilitate their accessibility and their usability. However, there are many contradictions between their goals in term of economic growth and the socio-technical process of opening up public data. Before publication, PAs have to gather together and sometimes digitalize data from different departments or services. In this context, data have to be standardized according to a common referent. So data is not completely “raw” (Baker & Bowker, 2007; Edwards Mayernik, Batcheller, Bowker & Borgman, 2011). Data is formatted, extracted from its initial pattern of categorization (e.g. vocabulary specific to a service or/and a task etc...) and harmonized (e.g. by using a common unit of measure). Whereas PAs have the aim of boosting

employment by encouraging the development of sustainable enterprises, this data production process may restrict open data re-use to specific fields of activity, and so to a limited number of companies. Furthermore, some available data are not published because PAs prefer to keep private certain information that could pose a threat for the local economy, the environment and citizens' safety. More precisely, they express reticence about the opening of data that may call into question the economic attractiveness of their region (e.g. negative effect of atmospheric pollution on property value) and data that may lead lobby groups to publicly express criticisms or shortcomings in certain fields, such as the adaptation of public transport to persons with disabilities. This selective publication process makes some data confidential and does not allow companies to create all the useful and operable tools that they could. Finally, on the current platforms on open data, PAs regulate exchanges. So, they can remove any information that would be unfavourable for them (e.g: an entrepreneur who develops a system which gives a solution to a specific claimed shortcoming with respect to local politics). Consequently, interaction is not completely free. In certain circumstances, entrepreneurs would not be able to build collaboration and so their activities could affect economic growth.

5.1.5 WISE&MUNRO/DEN HAAG

The case in Den Haag focuses on two groups of participants; PA (department of social affairs and employment) and employers. Collaboration between those two groups is common, but – as both groups agree - not always lead to new or creative ideas in an area with challenges for which out of the box thinking is needed. Reason for the municipality to get involved in the Route-to-PA project was therefore to open up the established forms of collaboration, and in combination with the focus on open data, integrate a new (online) way of co-creation in their working practice.

The preparation work in Den Haag in the first year of the project focused on the group, collaboration, open data and the tool. We worked on these issues in face-to-face meetings, sometimes with (individual) assignments in between the meetings. The group consists of four employers and three PA's to start with, and from the second meeting on, two open data specialists from the municipality and someone from smart cities The Hague also joined the group. One could say that this group consists of the main actors in this area of the municipality.

Besides these activities, we also contacted other main players in the field of open data to bring the project under attention. The group is considered to be a group of forerunners, paving the way for this new way of working.

Meeting 1 – October 8th, 2015

As researchers we were in charge of the program of these meetings. We suggested the topics to discuss and how to do that. We used methods (open questions; asking the other side; polite counters: yes, but...) that evoked discussions between group members, also because we wanted to see in the meetings how the participants are used to discuss among each other. We did not intervene much in these discussions, other than summarizing or keeping it on topic. We recorded the meetings for later analysis (camera and audio, see below), with the consent of the participants.

To set the goals of the meetings we painted the picture in four prospects.

1. We understand more about Open Data

By showing some examples of (relevant) open data and where to find these data, we explained the goal of the project in using open data in discussions between PA and employers. We noticed that for the employers in our group the idea of open data being of use was not new, and they were quite enthusiastic about it. For the PA's however, the concept of open data was more at a distance, probably because the department of social affairs

and employment is not yet part of any open data initiative of the municipality. Their ideas about opening up their data always brings up the issue (barrier) of privacy.

2. We are discussing something relevant and useful

We explained about various goals of the collaboration and what to put in and get out of each collaborative situation. Employers and PA instantly focused on co-creation as where they would like to go. They talked about the need to be innovative in topics like unemployment and also about sharing good practices when they are there. They seem to really want to open up the closed world of collaboration between PA and employers and use more transparency to learn from each other.

3. We are using a tool that supports us

We showed some pictures of SPOD from Vittorio's PowerPoint and explained a bit about the social medium and the role of open data. There was no version of SPOD available at that time, so we talked about the tool in general.

4. We are participating in a pilot in 2016 and preparing for this in these meetings

This is where we introduced the pilot with the alpha version of SPOD, and explained what was expected from the participants in 2016. We showed a planning for the next 6 months.

To get more understanding of the current situation of the users and to start the search of a common theme for the pilot in 2016, we asked the participants to share a narrative of their latest collaborative encounter with each other. What followed was a lively discussion, in which many characteristics of the collaboration between the PA and employers were seen and many themes were talked about. We summarized the discussion by pointing to some common points, themes that were mentioned by both employers and PA, and highlighted their wish to engage in a more meaningful conversation with one another, in which boundaries could be crossed and of which results would be shared among other interested parties. We ended the meeting with some concluding remarks and a short overview of what we would do in the next meeting.

Meeting 2 – November 9th, 2015

This meeting has as a main goal to know more about open data and open data policy of the municipality. For this reason, we invited the open data manager to participate in the meetings and pilot and asked him to give a short presentation on open data policy of the municipality. With this presentation all participants learned about the way the municipality experimented with opening up the data, how the national policy on open data was concretized locally. It became clear that in the department of social affairs and employment opening up data had not been a priority, but that the PA were thinking hard on how to make all decisions and regulations more open/transparent for all actors involved. We wanted to get the participants thinking more about how open data could be relevant for co-creation and creative problem solving in the domain of employment, so we

Example Case

Individuals who are older than 50 and are unfortunate enough to lose their jobs remain without a job for longer periods than before. As a consequence, their distance to the labour market is increasing. Would it be possible for city council and employers to conceive of solutions to this problem? What kinds of Open Data can be useful here? For example in order to:

- * Gain more insight into the nature of the problem
- * Support creative solutions

Where do these data come from, who owns them, where will they be shared, and what relevant information to you possess yourself?

We then left the topic of open data and discussed the different positions of the employers and the public administrators. Employers were supposed to have clear interests and were allowed to talk more freely. PA's were bound to higher order regulations and it was unclear to what extent they would be able to commit. Clear boundaries between private and professional opinions, and support from higher levels would help an open discussion. Increased mutual understanding between both parties is one of the outcomes of the whole project.

Meeting 3 – January 7th 2016

We started with summarising the previous meeting, and briefly introducing the project for a new participant in our meeting, from the municipal open data department. We explained about the current status of tool development and showed some slides of part of SPOD as a teaser.

Online discussion: experience and how to do it?

The next part was about online discussions, what was everyone's professional experience with such activity, and how did they feel about that? The consensus was that people do not simply participate in any forum with anyone. We need a clear motivation and goals. Would a professional platform be useful to achieve more depth in discussions between different organisations, such as WSP (employers service point, the municipality) and employers? The current situation is that yearly face to face meetings are pleasant and useful, but do not produce many new ideas. With the picture below we showed an abstract representation of a discussion, with open data in several roles: elaborating an idea, providing counterevidence, presenting an alternative, answering a question or even supporting a conclusion.

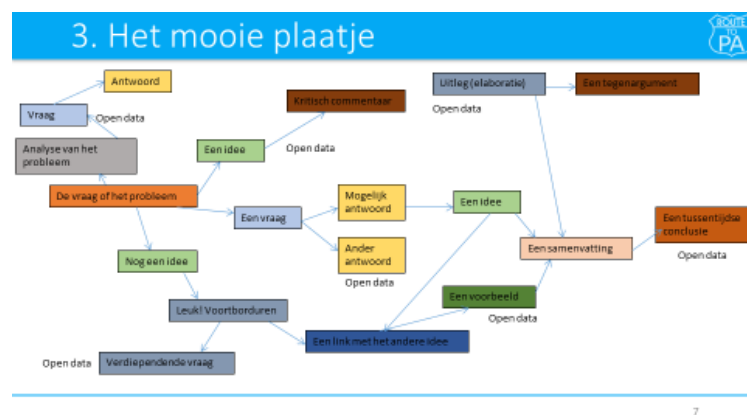


Fig. 6 Example of an online discussion with open data

We presented the characteristics of good interpersonal discussion: common goal, cognitive and ethical consideration and development of trust (Allwood, 1996). As simple as this may sound, it is very difficult to achieve and not many discussions score high on (apparent) consideration.

Transparency

From the discussion many thresholds came to light for complete transparency to be achieved. Trust within a group only develops in small rooms, working on a limited task. The issue of backrooms was taken up, and people agreed that transparency was not the same as simply opening up everything. It could also be transparent to announce that government and employers have reached agreement on something, without revealing the complete process.. A more public debate, however, can sometimes prevent stakeholders to simply

repeat their own interests. A more closed debate would be better for developing trust. People talk more freely if they know everything will be erased afterwards.

Themes and question for discussion

We ended the session with some homework for the participants: could you come up with some suitable questions for the discussions we plan with the tool prototype alpha?

Next steps were discussed. The group stays involved in the next year by means of several pilots; small usability studies and actual discussions.

6 EVALUATION AND TRANSPARENCY

6.1 VIEWS ON TRANSPARENCY

One of the main goals of the project is the evaluation of the tools. The purpose of the evaluation is to establish if various ways of working with our tools at the sites, in combination with other activities of the project has resulted in changes in government transparency. In our vision, this also implies the symmetrical change in view by the government towards its citizens: the citizens have become more transparent for local government. For both parties the same maxim would apply, that this transparency will not be complete, but only to the extent that collaboration can rely to a greater extent on mutual respect and understanding. Specific for local government is that civilians have a greater sense of influence and control. From their side, local governments have greater insight in what their civilians want and need from them. All of this is covered by the simple concept of transparency. Therefore, an introduction is needed.

Transparency is a concept that is used in different disciplines and that can have a different meaning. Building upon the work in WP 3.1, the Societal Activity Model can be used to distinguish the following views on transparency:

- *Technological transparency.* This view highlights that transparency is created when information is available to be used. In the literature this is sometimes referred to as nominal transparency (Heald, 2006: 34). This nominal transparency is generally assessed in terms of number of documents and datasets that are made available.
- *User transparency.* This view highlights that there must be receptors capable of processing, digesting and using the information. Heald (2006: 35) refers to this type of transparency as effective transparency. Effective transparency is assessed in terms of the accessibility, understandability and usefulness of the information to users.
- *Community transparency.* This view stresses that transparency is about social information: the meaning that is attributed to information in social processes. Transparency is constructed socially by group interactions and learning. Community transparency can be assessed in terms of the 'closure' that takes place within a group concerning the meaning of certain information.

- *Institutional transparency.* This view stresses that transparency is generated if more openness about performance, processes, roles, tasks, responsibilities, etc. has been created. Transparency acquires an institutional meaning when it influences interactions in the public domain. Societal transparency is assessed in terms of the institutional openness that is generated by interactions in the public domain.

These different views on transparency build upon another – technological transparency may generate user transparency that in group interactions produces community interactions and these communities may produce institutional transparency – but the relations are not straightforward. There may be barriers at each level and each level may put a different emphasis on certain forms of transparency. This means that transparency needs to be evaluated from different perspectives to produce a holistic understanding of SPOD and TET.

6.2 SCENARIO FOR THE EVALUATION PLAN FOR YEAR 2

6.2.1 OBJECT OF EVALUATION: ACTIVITY MODEL

Setting up an evaluation methodology for this type of socio-technical intervention with implications at the individual, group and societal level is highly complicated. To evaluate ROUTETOPA, we need to be specific about the objectives to be realized and the dynamics that are taking place at the different levels. Following up on WP3, an Activity Model Evaluation is needed: the different elements of the model can be translated to transparent at the level of the technology, the user, the community of users and society.

We formulate the following objectives for ROUTETOPA at these four levels:

Table 1. Activity Model Evaluation

Level	View on transparency	Activity Model	Focus	Objective
TECHNOLOGY	TECHNOLOGICAL TRANSPARENCY	TOOL	FUNCTIONALITY	TET and SPOD provide the functionality for enhancing Public Sector Transparency.
USER	USER TRANSPARENCY	SUBJECT	USER EXPERIENCE	Users experience a higher Public Sector Transparency because of the use of SPOD and TED.
COMMUNITY	COMMUNITY TRANSPARENCY	COMMUNITY, RULES & ROLES	COMMUNITY INTERACTION	The (digital) communities generate a higher Public Sector Transparency for the participants.
SOCIETY	INSTITUTIONAL TRANSPARENCY	POLICY ISSUE, OUTCOMES	VALUE FOR PUBLIC DOMAIN	The (digital) communities produce a higher Public Sector Transparency for society.

The following issues are of specific importance:

- The variation between the cases will increase when we move from the lower to the higher levels: the technology can be expected to function in the same way in each pilot, user evaluations will show some variation, group dynamics will be even more different and societal implications will vary most.
- The users are both citizens and government officials. They engage in co-productive processes of learning about issues, problems and solutions. This means that we need to investigate user experience

both among citizens and government officials (and probably also other users such as experts and stakeholders).

- We realize that transparency is instrumental to democratic processes but the realization of a stronger democracy is beyond the reach of this project. We evaluate the instrumental value of more public sector transparency.
- The levels are studied independently but also in relation to each other. The dynamics between the levels may prove to be most interesting. We can analyze, for example, how group dynamics (community level) influence public value (society level).
- Transparency has a different meaning at each of these levels: technical transparency, user transparency, group transparency and societal transparency. This means that we will have to develop a multidisciplinary definition of transparency. This can probably be a valuable addition to the literature.

Our proposal would be that the different research groups – Utrecht University, CNRS, Galway and Wise & Munro – will assume responsibility for these levels. They will develop a format for evaluation and this will be implemented at the different pilot sites.

6.2.2 FUNCTIONS OF EVALUATION: ASSESSMENT, DESIGN AND LOCAL DEVELOPMENT

Monitoring and evaluation fulfill a range of functions in the project. More specifically, M&E serves for:

- *Assessment or comparative benchmarking.* This means that we want to compare the different cases to see whether we see similarities or differences. This type of evaluation is needed to draw more generalizable conclusions about transparency.
- *Design of TET, SPOD and GUIDE.* This means that we need to draw lessons in each case about the usefulness of TET and SPOD and that each case needs to contribute to GUIDE by identifying good and bad practices.
- *Local case-specific development and learning.* This means that we will evaluate the specific dynamics of each case to see to what extent this case realizes the goals that have been set by the public administration and stakeholders.

This means that we will need a differentiated approach to M&E. To evaluate ROUTETOPA, we need to develop evaluation systems for comparative assessment, socio-technical design and for local development and learning. The three approaches are summarized in Table 2 below.

Table 2. Overview of different functions of evaluation

	M&E for Assessment	M&E for Design	M&E for Local Development
	Independent assessment for benchmarking and control	Participatory assessment for the design of technology	Shared assessment for situated learning
1. Who	Salerno: Technology Galway: Users	Salerno with the help of all cases	Prato Issy

	CNRS, WM: Community Utrecht: Institutional		Dublin The Hague Groningen
2. What	Predetermined indicators	Functionality of the system	Members identify their own indicators of success
3. How	Scientific objectivity	Combined socio-technical assessment	Self evaluation, method adapted to local culture
4. When	After program or mid-program	Frequently	More frequent small scale evaluations
5. Why	Accountability and control	Develop better technology	Empower community, corrective action

The following issues are of importance:

- These types of evaluation are not fully interdependent. They may partly overlap and they should certainly feed into each other.
- The role of the researchers is quite different in the different types of evaluation. Researchers are, respectively, experts, technology developers and moderators.
- Each research team develops part of the evaluation methods, and we are aiming for application of all general levels of study at all sites, albeit we do not expect this to be possible everywhere in an identical manner.
- The assessment of technological transparency will be developed by Salerno. It focuses on issues such as the stability of the system and the function that are provided. This evaluation is developed further in D4.1, under usability work.
- The assessment of user transparency will be developed by Galway. This analysis focuses on issues such as information needs, social interaction needs, understandability and usability needs. This evaluation is developed in section 5.3.
- The assessment of community transparency will be developed by CNRS and WM. This analysis focuses on issues such as number of participants, number of contacts, rules in the community, roles in the community, interaction patterns, appreciation, etc. This evaluation is further developed in section 5.2.
- The assessment of institutional transparency will be developed by UU. This analysis focuses on the contributions of transparency to monitorial, participatory and deliberative transparency. The evaluation is developed in section 5.1.

6.3 LEVELS OF EVALUATION: GENERAL PLANS

6.3.1 THE SOCIETY LEVEL

At the societal level we are trying to enhance our understanding of the public value or the outcome of the use of the tools on the stakeholder environment (other affected citizens, PA's, businesses, etc.). This implies the policy issues and desired outcomes at the organisational level: participants and their colleagues in the community may have changed their views about transparency, to be established by carefully phrased interview

questions. In general it is important to establish how people inside of the community view transparency, and how this is manifested in the meeting they frequent. Our participants could suggest possible candidates for interviewing and monitoring. Also at this level we suppose the satisfaction and the numbers of users are relevant. A very important criterion, especially for collaborative scenarios is the following: what happens with the outcomes of discussions? Finally, we can have a theoretical conception of the kind of democracy and the views of transparency that would fit with such a view. Do we see transformations?

The assessment of institutional transparency will be developed by Utrecht University. The analysis focuses on the perceived contribution to monitorial, and participatory democracy. The indicators are presented below:

Monitorial: contribution to insight in government performance / better checks & balances?

Perceived contribution of the community to critical assessment of government performance (by participants).

Perceived contribution of the community to critical assessment of government performance (by research experts)

Deliberative: contribution to public debates?

Perceived contribution of the community to public debate (by participants)

Perceived contribution of the community to public debate (by research experts)

Participatory: contribution to new forms of co-creation?

Perceived contribution of the community to new forms of collaborative action (by participants)

Perceived contribution of the community to new forms of collaborative action (by research experts)

General impact

References to the community in media reports.

References to the community in policy plans.

References to the community in council minutes.

References to the community in other virtual communities.

Discuss the societal impact in the SPOD.

A measurement instrument will be developed on the basis of these indicators. The institutional analysis will start in the spring of 2016 with a literature review to identify objective indicators in the literature. Following a combination of participant interviews, expert interviews, non-involved stakeholder interviewing and media analysis. A specific protocol will be developed by UU and will be applied by the teams in the five case to generated information for a comparative analysis at the institutional level. Measurement will take place at the end of year two and at the end of year three.

6.3.2 THE COMMUNITY LEVEL

A major aim of the ROUTE-TO-PA project is to stimulate the **emergence and evolution** of productive **online epistemic communities of practice ("OECOP")**, around specific Public Administrations associated to the project. mediated by the SPOD-TET tools. On the level of such communities, therefore workpackage 5 is oriented towards determining the extent to which this aim is achieved.

The main dimensions of evaluation on this level can be formulated in terms of the characteristics of the **participants**, the **structure of their interactions** and their **content** within the OECOP, studied from a **developmental perspective** (Engeström, 1987). The developmental, or *diachronic*, analysis draws on *synchronic* analysis of the OE-CoP (at a given point in time), and identifies relevant differences across time-points, with a view to identifying the overall *trajectory* (Dreier, 1997; Ludvigsen et al., 2011) of the community, i.e. where it is heading.

- The focus on **participants** of the OECOP asks who the participants are (their professional status), their times of entry (and exit) from the community, their position in the social network (other persons with whom they regularly interact) and their interactive roles (i.e. the types of interactions in which they distinctively engage: see Détienne et al., 2015). From the developmental perspective, changes in participants' positions in the network, their evolving interactive roles, are tracked over the lifetime of the OE-CoP.
- The focus on **structure** of the OE-CoP aims to apprehend who is communicating with whom, how is communication structured into threads, what specific "zones of collaboration" emerge? The evolution of such structures over time is studied from the developmental perspective.
- The focus on **content**² of activity in the OE-CoP (messages sent using the SPOD, open data accessed, visualisations created and displayed) aims to understand the objects of activity (according to the Societal Model of WP3 Year 1), what are the participants joint aims, to what extent are new joint purposes (e.g. of open data provision) emerging, to what extent do the participants understand the open data visualisations? To what extent do the participants manifest self-efficacy? What are the participants' visions of changes in the degree of transparency of the PA, as made manifest by their activity in the SPOD-mediated OE-CoP? From a developmental perspective, how do contents focussed on evolve over time, and what evidence is there that the "tensions" identified in WP2, between young entrepreneurs and the PA, tend to diminish?

The main object of study and the dimensions of analysis are shown in the diagram below:

² We use the everyday term "content" here with reference to "content analysis" to refer to both the *semantic contents* of messages sent using the SPOD (e.g. specific elements of open data; open data visualisations; the SPOD too; participants' competences; the previous dialogue, ...) as well as their *communicative functions* (e.g. request, inform, question, critique, regulation proposal, ...). A better term might be "meaning".

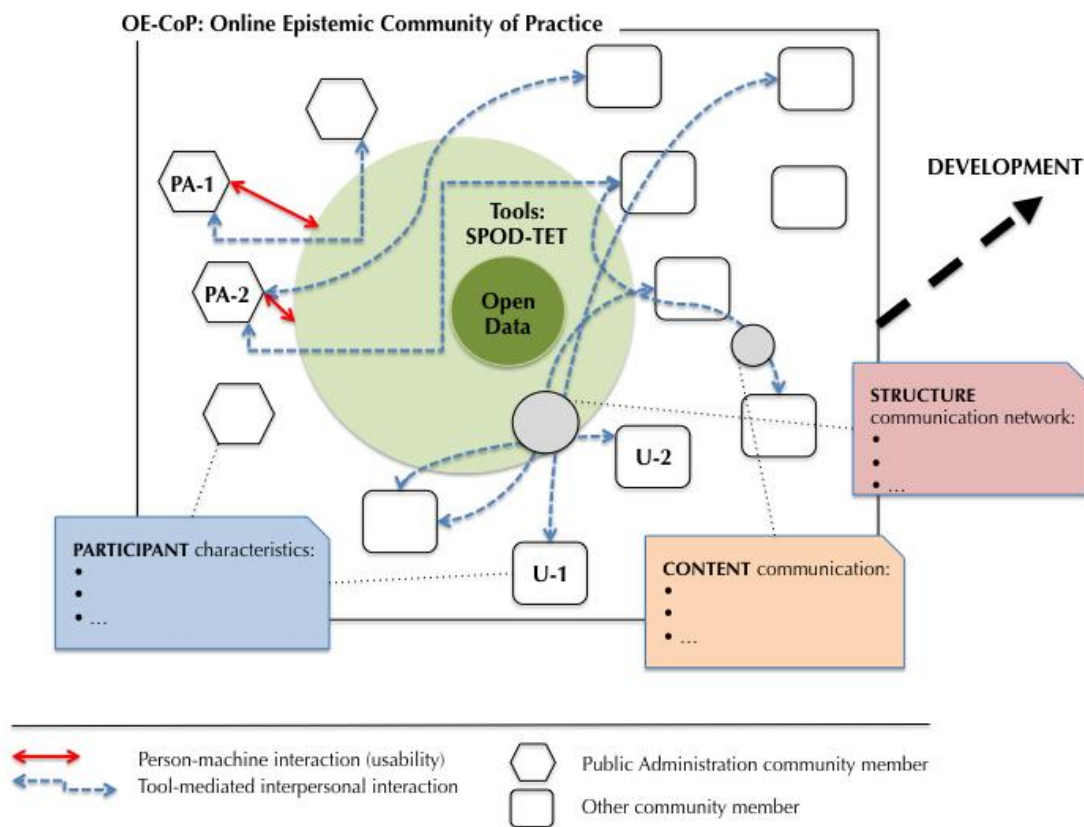


Figure 7: main object of evaluation (OECop focussed on open data), with three dimensions of analysis.

Communication on online forums can often take the form of *decoupled action*, where many users post a few messages then remain silent (either they leave the community, or remain as ‘lurkers’). In some cases, more tightly coupled *zones of collaboration* will emerge (Fréard et al., 2012: see figure 8 below), and these are signs of ‘communityness’. Clearly, the latter is to be preferred. Zones of collaboration (Baker, 2015) involve a *shared focus* (e.g. all of a group of participants focussed on the same open data/visualisation) and quasi-synchronous interpersonal communication involving all that displays effort towards shared understanding (Clark & Schaefer, 1989). Methodological tools have been developed for analysing the **quality of collaboration**, in on/off-line communities or groups (Détienne, Baker & Burkhardt, 2012).

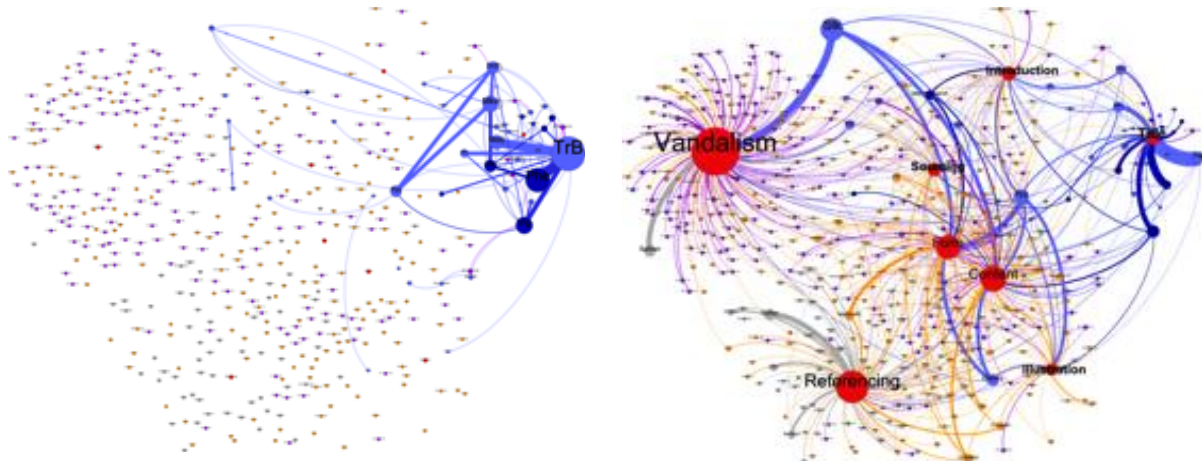


Figure 8: Visualisation of zones of collaboration in decoupled action (Fréard et al., 2010) in an online community (Wikipedia). Points represent messages in discussion space, thickness of lines is proportional to density of communication

The **general question** to be answered along the “community” dimension of evaluation is thus: **to what extent does the collective user activity mediated by the SPOD-TET tools correspond to and evolve towards that of a veritable online epistemic community of practice?**

A veritable online epistemic community of practice will have at least the following characteristics, to a high and increasing degree over time:

OE-CoP main characteristics:

- main **focus** on open data and visualisations (i.e. rather than managing interpersonal relations, talking about something else, or, crucially, focus on understanding how to use the SPOD-TET tools themselves);
- emerging **zones of collaboration** (focussed on shared contents, high density quasi-synchronous communication displaying effort towards shared meanings, involving co-creation of open data visualisations, ...);
- a variety of **forms of collaboration** (more or less symmetrical, involving varied and exchanging roles);
- emerging and reciprocal **interactive roles** (e.g. “regulator”, “critic”, “guardian of community rules”);
- progressive elaboration of **community rules**, or “norming” (Miller, 2003);
- a sense of **‘belonging’** to a community, manifested by participants;
- **supportive**, ‘cooperative’ behaviours (encouragements, helping, ...);
- progressive **integration** of legitimate peripheral participants (new participants progressively integrated, in terms of growing number of communications between them and other community members).

Research questions are studied both **synchronically** (state of OE-CoP at a given time) and **diachronically** or **developmentally** (evolution of states). We study the extent to which the OECoP evolves in positive terms with respect to the above characteristics.

These dimensions of analysis will be studied from two points of view: the **first person perspective** (the participants’ own views, representations, elicited, for example by interviews) and the **third person perspective** (analysis of the participants’ activity in the OE-CoP). Analysis of the OE-CoP is carried out **qualitatively** (by the

researcher, concerning, for example, communicative functions and semantic contents), where frequencies and distribution of analytical categories are calculated quantitatively, as well as **automatically** (numbers of messages per participant, lexical analysis, Social Network Analysis concerning thread/network structure).

We hypothesise that an OE-CoP that displays the above-mentioned characteristics — i.e. those of a collaborative, cooperative online epistemic community — will display processes closely linked to **a more transparent interrelationship** between citizen users and public administration users.

This work to be carried out within **Workpackage 5** also relates closely to the Models and Methods developed, over three years, in **Workpackage 3**:

- The **Societal Model** (year 1) relates to the above in terms of analysis of the object of activity (including community rules), and the expression of tensions. In these cases, the level of **content** analysis is relevant, for identifying ‘what the participants are trying to achieve’ (i.e. object of activity) in given exchanges, and, on the level of communicative functions, given that tensions will be correspond to argumentative functions and, most likely, to the salient expression of strong emotions.
- The **Community Model** (year 2) provides the theoretical background and methodological tools for addressing the main research question, concerning evolution of SPOD-TET mediated collective activity towards a veritable epistemic community.
- The **Social Representations Model** (year 3) concerns the evolutions of the OE-CoP participants’ **representations** (attitudes, appraisals) of the community, of their perceived self-efficacy, of the degree of transparency and cooperativity of the Public Administration, and of the SPOD-TET tools themselves. This will be studied using interviews and questionnaires, but also on the basis of **content analysis** (what are the attitudes expressed using SPOD?). This can draw on **appraisal theory** (White, 2011), i.e. the positive and negative attitudes and affects expressed in language.

Our main research questions, formulated in terms of analysis of relations between characteristics of **participants**, the **structure** of communication and the **content** of communication, give rise to the **evaluation and success criteria** described in the table below.

A “successful” collective SPOD-TET mediated activity involves close collaboration on open data and visualisations of it, effective regulation of the community and growth through integration of peripherally participating users: operationally, it corresponds to *a tendency towards increase in terms of the criteria described below*.

Specific evaluation criteria include the following:

Analytical dimension(s)	Criterion	Success/methods/remarks
Participants	N participants	Increase in N participants alone is not relevant. What is more important is the rate of growth of N participants, whilst preserving the community (collaboration zones).
	N participants × N messages × Average length messages	A measure of degree of participation and degree of (s)ymmetry of it. Automatic social network analysis. Success = increase in number of active participants (decrease in “free riders” or “lurkers”), increased symmetry of participation.
	Representation of transparency	Interviews, questionnaires, qualitative network analysis

		(analysis of appraisal of PA, tools, open data). Success = move towards more positive attitudes.
	Representation of self-efficacy	Interviews, questionnaires, qualitative network analysis
Content	Object of activity (open data understanding, community rules, making new contacts, requesting new open data, ...)	Success = migration towards more focus on open data use and understanding. Content analysis of messages, threads, interviews (what were participants trying to achieve?)
	Communicative functions of messages (e.g. informative, argumentative, explanatory, regulative).	Success = decrease in proportion of regulative activity, successful constructive resolution of argumentative conflicts, successful (mutually accepted) explanation and information sharing.
	Open data sets accessed	N (different) data sets accessed
	Visualisations created	N visualisations created × N data sets; N modifications to them (and N interface actions to implement modifications)
Structure	Degree of connectedness of social network	Analysis of threads, who communicates with whom; automatic Social Network Analysis. Success = increased connectedness, length of threads, segmentation
	Number and average length of threads	idem
	Segmentation into cliques, clustering	Idem
Participants — Structure	Position of participants in structure (e.g. initiator of thread, number of connexions to others, ...)	Success = change in position of participant in social network: peripheral legitimate participation → full participation, well linked into network
	Homophily (extent to which participants linked to those of same or different type (e.g. connectedness between entrepreneurs or between	Success = greater degree of communication between PA representatives and other types of users
	Mutuality/Reciprocity	The extent to which two actors reciprocate each other's interaction.
Participants — Structure — Content	Emergence and co-evolution of interactive roles	Interactive roles are clusters of communicative acts/semantic contents that characterise participants and differentiate them from others (e.g. critic, regulator, explainer, ...). In a genuine community, there is co-evolution and reciprocity of such roles — e.g. when one participant regulates the discussion less, others 'fill in' for this role. Success = increased differentiation of roles and reciprocity (signs of a 'mature' CoP)
	Zones of Collaboration	Extent to which cliques emerge (see above) that involve intense interaction, approaching synchronicity, on a common focus.
	Quality of Collaboration	Degree of mutual understanding, fluidity, constructive conflict resolution, task focus, etc. (Détienne, Baker & Burkhardt, 2012)

Table 6: Evaluation and success criteria at the community level

6.3.3 ANALYSIS AT THE USER LEVEL

The Evaluation Model - Technology Acceptance Model

To evaluate the artifacts, we shall develop a framework based on the well-known and studied Technology Acceptance Model (TAM). The model was published by Davis (1989) and is the most widely accepted model for

understanding the usage of Information Systems (IS) and its acceptance. It suggests, that external variables (such as system design and rich features) influence the perceived usefulness and perceived ease of use. Over time the model evolved to TAM2 (Venkatesh and Davis, 2000) and was extended with additional external variables, relevant to IS utilised in the workplace: the social influence variables (i.e. subjective norm) and the cognitive instrumental variables (i.e. result demonstrability). Original TAM is presented in Figure 9 while TAM2 is shown in Figure 10.

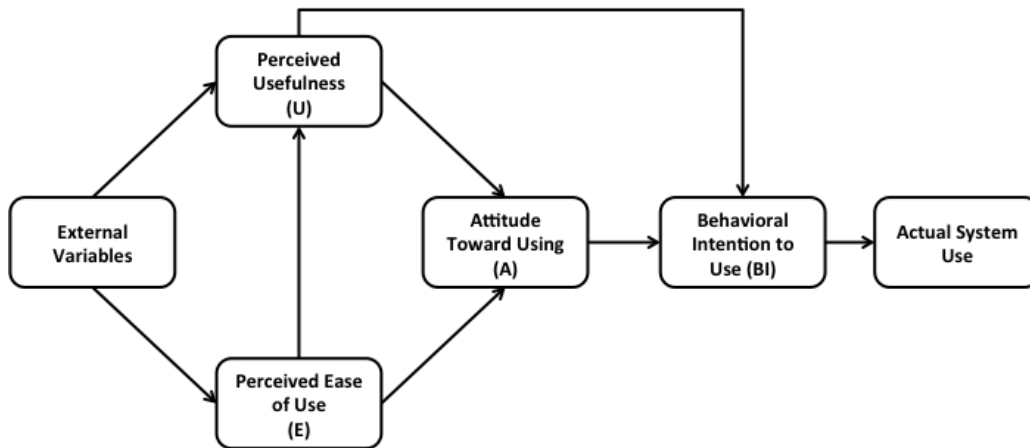


Figure 9. The Technology Acceptance Model (Davis et al, 1989)

The definitions of the additional variables in TAM2 are defined as follow:

- **Voluntariness** - *“the extent to which potential adopters perceive the adoption decision to be non-mandatory”.*
- **Subjective norm** - *“a person’s perception that most people who are important to him think he should or should not perform the behavior in question”*
- **Image** - *“the degree to which use of an innovation is perceived to enhance one’s status in one’s social system”*
- **Job relevance** - *“an individual’s perception regarding the degree to which the target system is applicable to his or her job”*
- **Output quality** - *“the tasks a system is capable of performing and the degree to which those tasks match the job goals”*
- **Result demonstrability** - *“the tangibility of the results of using the innovation”*

Evaluation Criteria

Based on the Technology Acceptance Model in the pilot evaluation, we evaluate the following parameters:

- **Relevance** – how relevant is the use of open data and the TET/SPOD to the user’s job and daily life,
- **Output quality** – what is the quality of datasets available on the platform? What is the perceived quality of TET and SPOD platforms?
- **Result Demonstrability** – does the use of the TET and SPOD address the relevant user needs provided by users during the Scenario development workshop?
- **Perceived Ease of Use** – how easy is it for non-technical users to use TET and SPOD?
- **Perceived Usefulness** – how useful do users perceive TET and SPOD?
- **Intention to use** – how willing are the users to use TET and SPOD to support their information needs and decision making needs?

The criteria has been integrated into detailed User Evaluation Guide as questions to completed after carrying a set of tasks associated with the implemented platform features.

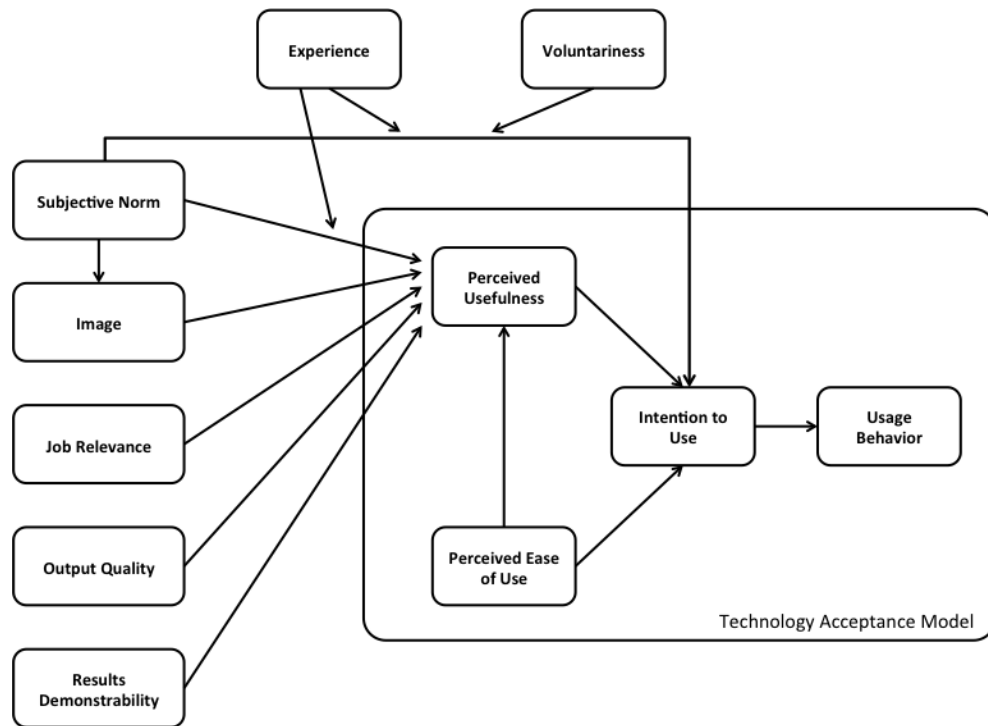


Figure 10. Technology Acceptance Model 2, (Venkatesh and Davis, 2000)

Data Gathering Strategy

Evaluation data will be gathered by organizing the focus group. The following methods will be used:

- *questionnaire survey* – the survey will be used to collect the information from the individual users in the form of structured questionnaire with open and closed questions. The information will be collected with the use of unified scale in order to enable comparison.
- *interview* – the interviews will involve discussion: talking and listening of individual users. It some cases it may follow previously defined structure.
- *scenario testing* – Scenarios will require users to follow step-by-step guide and to provide feedback on his/her experience with the system.

Method	Advantages	Disadvantages
<i>questionnaire survey</i>	<ul style="list-style-type: none"> • Easy to create • Enables comparison of results • Enables statistical analysis 	<ul style="list-style-type: none"> • Low level of details • May not easily identify specific issues
<i>interview</i>	<ul style="list-style-type: none"> • Flexible level of detail 	<ul style="list-style-type: none"> • Difficult to compare

<i>scenario testing</i>	<ul style="list-style-type: none"> • Easy to create • Context specific 	<ul style="list-style-type: none"> • May focus too much on the scenario not on the system as a whole
-------------------------	--	---

Each of mentioned methods has its advantages and disadvantages (as shown in the table above), which is the reason to use mixed methods' approach during the evaluation process.

This instrument will be used to collect evaluation data about the both TET and SPOD (See Section 6 of Deliverable 4.2). The aim will be to collect quantitative and qualitative data. Analysis of the gathered results will be used as the input for refining both TET and SPOD subsystems comprising the ROUTE-TO-PA platform. Apart from the evaluation, detailed both TET and SPOD platform will be tested from a technical point of view. See the Appendix of Deliverable 4.2 for details of the test specifications.

For each evaluation cycle, two levels (User and Group levels) of results evaluation will be carried out. The evaluation cycle correspond to the end of an official release of the TET and SPOD platforms. For instance, the first evaluation of the alpha versions of TET and SPOD will be carried out soon after their release at 31 Jan 2016.

7 RESEARCH PLANS FOR YEAR 2

7.1 PRATO

The site of Prato is not a specifically linked to a research team. Therefore, evaluation activities will be at the technology, society and user levels. Particular research plans will be further developed after the workshops in February.

In the second year of the project the pilot activities will start first by involving a restricted number of users, in order to collect feedback and check how the usage of the SPOD/TET platform is adequate to users' expectations. Secondly, the pilot activities will be opened to a wider audience with the help of the project local Point of Presence (PoP), where local stakeholders will be involved to keep discussions alive and productive. During this phase, appropriate monitoring and evaluation tools will be used, on the basis of the methodological approach developed in the project but also specifically tailored for local needs.

The following table (Table 1) reports the planned scheduled activities for year 2 in Prato.

	1	2	3
Date/Period	11 February 2016	15-29 February 2016	April 2016 -January 2018
Title	Phase 1: Workshop	Phase 2: Activity at home	Phase 3: open experiment
Description, aim	At least 10 stakeholders will participate in this first workshop, where they will be asked to use the SPOD/TET platform on the	The participants in workshop 1 will be asked to keep on testing the SPOD/TET platform and to build discussions on the city	In this phase a wider city audience will be involved in the pilot and the project opportunity will be advertised through several

	<p>basis of a pre-defined scenario</p> <p>The session will last 2h30.</p> <p>The goal is to verify the impact of the platform and the GUI usability, but also the accessibility and comprehension of open data and the interaction among users.</p> <p>The results of the workshop will represent an important feedback for the planning of the following pilot activity with a wider audience.</p>	<p>budget issues in their spare time. This will represent a simulation of the usual usage of the platform.</p> <p>At the end of the period (that might be extended furtherly if necessary), participants will be interviewed in order to collect their final impressions and to gather ideas for the planning of the following pilot activity. The interviews will also allow to get some information useful for the modelling of the data-based social interaction through the SPOD/TET platform.</p>	<p>channels (papers, newsletters, leaflets, etc.).</p> <p>The focus will be at first on the city budget and will include also the verification of expenditures as long as the city administration spends the budget.</p> <p>Later on, we expect the SPOD/TET platform be used also to discuss on different issues of interest for the citizens (e.g. immigration) and the effort will be that of adding always more datasets in the local CKAN platform, to support discussions, but also to encourage the access to other datasets elsewhere available through the SPOD/TET platform.</p> <p>Some monitoring and evaluation tools will be employed to analyse the current usage of the platform and to identify interesting elements for social interaction models and the project final guidelines.</p>
Participants	<p>Businesses, PAs and researchers. Some of them were already involved in the workshop for the identification of user requirements (held in Prato on 23 April 2015) and in the presentation of the SPOD platform on 28 September 2015.</p>	<p>Same as in workshop</p>	<p>Citizens and stakeholders (businesses, students, associations, etc.)</p>

Table 1: Planned pilot activity in Prato in the second year of the project.

7.2 GRONINGEN

Summary of Research Approach (UU/Groningen)

Research Purpose: 1) Enhancing the understanding of the public value of the re-use of open data among pa's and citizens groups in dialogue and collaboration (*Society*). 2) Facilitating situated learning in the community in discussion and collaboration regarding population decline issues (*Community*) 3) Designing and testing tools for open data use in relation to discussing and collaborating regarding population decline issues (*User and Technology*)

General research question: How can the re-use of open data by pa's and citizen groups contribute to dialogue and collaboration between citizens and pa's regarding population decline issues? More specific questions are: a) How do TET and SPOD contribute to dialogue and collaboration between citizens and pa's regarding population decline issues? b) How do users perceive the contribution of use of open data via TET and SPOD, to dialogue and collaboration between citizens and pa's regarding population decline issues? c) How do the communities on TET and SPOD interact and collaborate regarding population decline issues? and d) What public value does the use of open data on TET and SPOD in terms of participation and collaboration for population decline issues?

Data context: A longitudinal case study with a scenario-design based intervention will be conducted in the Province of Groningen, from March 2015 until February 2018. The case study will consist of three functions of evaluation: Monitoring and Evaluation for assessment, Monitoring and Evaluation for Design and Monitoring and Evaluation for Local Development

Data sources: content analysis, a questionnaire and expert judgement; usability tests of TET and SPOD: expert testing, student testing and panel testing, based on qualitative methods e.g. researcher constructed narratives; for Local Development, a project team in Groningen of users and pa's will provide input for shared assessment and situated learning (Bryman, 2012) e.g. interviews, focusgroups, participant observations, document analysis and scenario's.

Data analysis techniques: qualitative and quantitative analysis; content analysis user feedback on TET and SPOD and language based approaches will be studied to collect qualitative data such as discourse analysis (Bryman, 2012).

Unit of analysis: Individual users (targeted intermediaries, followed by a focus on citizens in general); PA users (initially the focus is on pa's at the Province of Groningen, followed by a focus other pa's working at municipalities); the local community: e.g. communities around citizens initiatives in Province of Groningen; c) the society: both the perceived contribution of the technology to public sector transparency for society and on the other hand e.g. mentioning in council meetings or media

Underlying Societal Activity System: deliberative democracy with participatory elements.

Outcome: to facilitate situated learning in the community in discussion and collaboration regarding population decline issues

View of Transparency: Transparency and trust can be seen as an enabler for open participation and collaboration (Lee & Kwak, 2012). However transparency can also be the outcome of participation instead of a precondition (Wijnhoven, Ehrenhard, & Kuhn, 2015) which might indicate a dynamic mutually reinforcing relationship.

In Groningen TET and SPOD are tested by both students and actual users. We distinguish three phases. The alpha version is a preliminary version of TET/SPOD. In order to avoid project fatigue researchers, pa users and citizen users will provide feedback on the alpha version. In the second year of the project the beta version is tested in an experimental setting by students and finally in the third phase in a natural setting by users involved

in citizens initiatives in the Province of Groningen. Next to testing this last phase is also focused on “community building” among the participants. The following testing activities are foreseen:

Phase 1 Expert Panel Testing TET/SPOD Oct 2015 – Feb 2016: In the first phase between October and January the researchers provided feedback on TET and SPOD based on knowledge obtained from the workshops (Work package 2) and the Societal Activity model (Work package 3.1.). As of February 2016 the Groningen project team will also start providing feedback on TET and SPOD. In this initial phase, where the tools have limited functions, the feedback is left open-ended, focused on what users (4 researchers and 8 users) encounter when using the tool.

Phase 2 Student Panel March - April 2016: Next to what people think they do, it is also investigated what they are doing by analyzing their activity with the tool. This will be done based on a “*Population Decline Challenge*” among students. In an experimental setting, participants will be asked to find solutions for the issue of population decline. In order to find solutions they can use the available data on TET/SPOD but also suggest other datasets that they think might be useful. Furthermore, it is important that they exchange ideas (deliberation) on the platform and work together (collaborate), using each other’s expertise in order to find solutions and innovative ideas. Finally, they should also be able to contact a PA for questions or requests for more information. We will be testing TET/SPOD at the four different levels; Technology, User experience, Community but also Society in terms of how the participants perceive what the contribution is to public sector transparency. 10 bachelor students will be recruited. Each student will be asked to provide input and participate in discussions with each release. The students willing to participate will be invited to join a start session in which the project and what is expected from them will be explained.

Phase 3 User Testing and Community Building

As of Spring 2016, when the beta version is ready, we will start with testing the tools among a broader range of users, within, but also outside the “Population decline project group”. It will be the start of community building, while at the same time working on visualizing, sharing and discussing data (see the phases during the working process)

7.3 DUBLIN

The pilot evaluation for the Dublin Pilot will draw upon the expertise of partners in relation to societal and community levels of analysis, but will primary be focused on driving and coordinating the user level analysis using the Technology Acceptance Model (TAM). The model measures “Ease of use” and “Perceived usefulness” of the TET and SPOD by the target users. In Dublin, a focus group partly comprising participants of the Collective Intelligence Workshop held in April 2015 in Dublin for developing Scenarios and User stories for TET and SPOD will be constituted as target users for evaluating both platforms. Evaluation goal is to check if the perception of users about TET and SPOD affordances cover the indicated social and interaction needs in addition to the understandability, usability and decision making needs that were provided earlier by users (see Deliverable 2.3). These were refined into use-cases and software requirement that guided the development of TET and SPOD. See Figure x for the process to be adopted in evaluating the TET and SPOD.

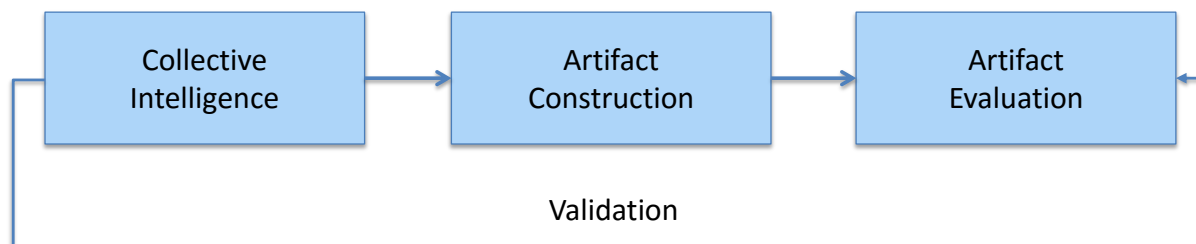


Figure x: Evaluation of TET and SPOD platforms

Planning of Evaluation Activities

A series of four workshops will be scheduled between March 2016 and January 2018. Exact dates are yet to be determined. The workshops will provide the participants in the focus group opportunities to interact with and use TET and SPOD for evaluation of the available features. The four workshops will aim to evaluate:

- 1) Alpha version of TET and SPOD at Month 12
- 2) Alpha version of SIM in Month 18
- 3) Beta Version of TET, SPOD and SIM in Month 24
- 4) Final Version of TET, SPOD and Sim in Month 36

This work will be complemented by a series of activities, including:

- Launch of the new Dublinked and Smart Dublin websites with integrated SPOD and TET functionality
- Ongoing issue based engagement with citizens and stakeholders to identify and refine challenges relevant to sourcing relevant open data and engaging in scenario-relevant community building activities using SPOD and TET
- *Youth* workshop and competition focused on use of open data and application of SPOD and TET tools to community building activity
- *Citizen* workshops and competitions linked to ongoing project activities engaged by *Your Dublin Your Voice*, Dublinked, and Smart Dublin.
- *Stakeholders* workshops and information days showcasing available open data to wider Dublinked community and citizen groups
- *Dublinked users* thematic events, developer days, dubmeets focused on developing usage scenarios for SPOD and TET
- Plan for challenge competitions in 2016 to engage business, entrepreneurs and citizens in the development of new ideas and solutions to identified issues

7.4 CNRS PARIS

Summary of research approach (CNRS/Issy-les-Moulineaux)

1. **Research Purpose:** to understand the characteristics and evolution of the online communities of practice mediated by the SPOD-TET tools.

2. **Research Question:** to what extent does the user activity mediated by the SPOD-TET tools correspond to and evolve towards a veritable online epistemic community?

Particular sub-questions are: a) Identification of zone of collaborations and zones of decoupled actions; b) Characterization of forms of collaboration and their quality; c) Co-evolution of forms of collaboration (analysed by a third person/researcher perspective) and self efficacy (first person/participants viewpoint); d) Relationship between emerging roles (e.g. boundary spanner) and knowledge co-construction (epistemic content and communicative functions of exchanges).

3. **Data Context:** The study will start in June 2016 and will last until the end of the project (and hopefully beyond). Two workshops (Sept 2016, June 2017) will be organized to support community building, to understand the motivations of participants and to collect feedback on SPOD/TET. Participants will be citizens/young entrepreneurs (and potentially PA) using SPOD/TET with OD on the site of Issy media.
4. **Data Sources:** Video recordings of the workshops, automatically recorded traces of actions in SPOD and TET (discussions, Open Data use), individual interviews, online questionnaires.
5. **Data Analysis techniques:** Content analysis, SNA (automatic structural analysis), analysis of discussion (qualitative -based on coding scheme- and quantitative analysis –role identification), distinction between dialogic functions and epistemic content of the exchanges.
6. **Units of analysis:** the community, the zone of closely coupled action (collaborative zone), the thread and exchange/sequence on a common topic/object, the communicative act (semantic content, communicative function). Level 1, or the goal-directed actions in small groups; Level 2, or the group level object-oriented activity
7. **Use of activity systems:** the Activity System model defines elements such as the “object” and “community rules” of activity that will be analysed as they are expressed in and emerge from interactions using SPOD-TET (content analysis).
8. **Underlying activity model:** Community of practices, (epistemic) online communities, dialogue models, collaboration models, argumentation models.
9. **View on Transparency/social representations:** Improved self efficacy, which reflects the evolution of individual and collective capacities in the use/sense making of Open Data and of their representations about the implications of their actions.

From January 2016 to January 2018, the SPOD/TET will be tested in three phases as showed in the table 2. Two categories of stakeholders will participate in these experiments: PAs from Paris region (Open data producers) and young entrepreneurs (Open data users) from the city of Issy-les-Moulineaux.

During the two first phases (March 2016 – May 2016), the CNRS team (J.Groff, F.Détienne, M.Baker) will introduce the platform to young entrepreneurs and PAs. They will engage them to use it collectively in order to evaluate the usability of the system. The interactive dynamics of the discussion will be analysed (see evaluation criteria above).

During the last phase (June 2016 – January 2018), a representative sample of each group (entrepreneurs, PAs) will manipulate the SPOD/TET during 18 months with the aim of developing the platform by uploading new data, encouraging the integration of new members and promoting the co-construction of new content.

In this way, the CNRS team will explore the development of the activity of each population and analyse the community building progress (Brown, 2001; Garrison, 2007) and the process of knowledge co-elaboration.

	1	2	3
Date (actual or proposed)	March 2016	May 2016	June 2016 -January 2018
Title	Phase 1: Workshop 1	Phase 2: Workshop 2	Phase 3: ecological experiment
Description, aim	At least 10 young entrepreneurs from Issy-les-moulineaux will participate to the first workshop. They will ask to evaluate SPOD/TET by basing on a usage scenario. The session will last 2h30 and will be recorded. The goal of this experiment is to study the impact of network-	At least 10 PAs from Paris Region will participate to the second workshop. They will ask to evaluate SPOD/TET. The session will last 2h30 and will be recorded. This experiments aims to study the commitment of PAs in the opening process of public data according to the requests formulated by young	At least 20 young entrepreneurs and PAs from Paris Region will participate to the phase 3. They will ask to use SPOD/TET during 18 months. Their interactions will be logged on a server and the CNRS team will interview them at least once a month. The CNRS research questions will be focused on the community building process and more precisely on the constitution and

	<p>centric tool on the cooperative design. The CNRS team will analyse: the impact of graphical interface on - the accessibility and comprehension - of open data, the information sharing process and the collaborative decision-making mechanism.</p> <p>Results of this workshop will be considered with the aim of highlighting entrepreneurs/users' needs in terms of information, interaction and tools.</p>	<p>entrepreneurs in the previous workshop. In this way, the CNRS team will explore PAs/producers' needs in terms of discussion spaces, integrative functionalities and support information. More precisely, the team will focus on their expectations concerning the transformation of open data in sustainable business with the aim of promoting accessible and friendly data format that could be re-usable by entrepreneurs.</p>	<p>development of an online epistemic community of practice, integrating users-young entrepreneurs and producers-public administrators.</p> <p>The team will carry out a synchronic and diachronic analysis of the participants, the content and the structure. They focus, not only on these entities individually, but also on the interactions between them with the aim of studying the evolution of the community composition, the quality of collaboration within communities and between them and finally their co-evolution.</p>
	<p>The overall goal of these two workshops is to study the socially activities of young entrepreneurs and PAs with the aim of defining an effective framework to design a socio-technical system that promotes the development of an online epistemic community of practice around the transformation of open data into applications and services.</p> <p>So, The CNRS team will focus on the analytical dimensions of characteristics of participants, the nature and organisation of their interventions within the online epistemic, with the perspective to engage these two populations to co-create communities devoted to the exploitation of Open data.</p>		
Participants	<p>Young entrepreneurs from incubator. People who are in the process of creating a company and to people who have created their company since less than two years</p>	<p>Public administrators' group from Paris Region (Issy and surrounding municipalities): representatives of geographic information system, representative of associative life (promote creation and development of local associations) and representatives of communication services (communication actions towards general public, media and partners cities)</p>	<p>Young entrepreneurs (Issy-les-Moulineaux) & Public Administrators from Paris Region</p>

Table 2: The principle phases of the CNRS evaluation of the SPOD/TET

7.5 DEN HAAG/WISE & MUNRO

<p align="center">Summary of research approach (Wise and Munro/Den Haag Municipality)</p> <p>Research Purpose: Exploiting Open Data for developing more effective forms of collaboration between WSP (Department of social services and employment) and citizens.</p> <p>Research question: To characterize the development of co-creation by civil servants and professional employers during the tool-supported process of discussing pertinent topics for collaboration with open data.</p> <p>Data Context: workshops during the full period, combined with scenario-based collaboration; scenarios are jointly developed. A longer term perspective is to include more members of the employer community. All activity is recorded.</p> <p>Data Sources: Video recordings, traces of tool use, open data used, interviews, narrative collection.</p> <p>Data analysis techniques: content analysis (appraisal) of interactions and narratives, quality of use (expert norms) and interpretation of open data; development of collaboration</p>

Units of analysis: collaboration and appreciation as they develop over time

Activity systems: allows to characterise tensions and evolution of the group objectives

Activity model: societal

Transparency: improved mutual understanding between local administration and employers, and possibly other stakeholders

Compared to other cases, den Haag municipality, and more in particular the dept. of Social Services and Employment, is less advanced in their working with open data, and their thinking about transparency. Moreover, there are differences between the political and administrative levels how this could be achieved. Therefore, we decided to adapt our ambitions on transparency to what is happening at both levels, independent of the project's achievements. During the first year, we discovered a high interest both at the level of employers in the municipality, and of some of the public officials we dealt with. As for the latter, we discovered interested and involved functionaries at the dept. of Transport. From there, initiatives toward a support unit for open data are ongoing. We expect some progress during the second year, which may allow us to involve other groups of stakeholders in our project. We think one should be very careful to call our work an intervention study, as the time for this intervention to make a permanent impact on the stakeholders involved is quite short. What we can hope for is to establish some small changes that might act as snowballs becoming bigger and more forceful.

Therefore, our long-term strategy is cyclical and design-based, with frequent evaluation and short periods of use. During the upcoming period our main focus will be on civil servants and professional employers. Their main interest, and hence ours too, is whether or not the platform is an improved area for them to collaborate and discuss mutually interesting topics. Our main research question therefore is formative: **To characterize the development of co-creation by civil servants and professional employers during the process of discussing pertinent topics for collaboration with open data.** This diachronic perspective involves tracing over a relatively long period of time how the group activities evolve, especially in terms of collaboration. What we are hoping for, and what we, as researchers, but including the participants as pioneers, are jointly trying to establish is the implementation of the advanced idea of collaboration as a main aspect of transparency.

Data will be collected, first during (parts of) face to face meetings that are held monthly, at a convenient time between 16:00 and 18:00 PM. Second, we will introduce the tool to the users, and they will start a pilot with the tool, as soon as the tool is ready and we have thoroughly investigated the possibility of good questions in combination with useful open data.

The first four face to face sessions will be compared to get an idea of the topics discussed, the progress achieved, and the state of the common ground for co-creation. The analyses will be qualitative and in first instance based on the outcomes of the society model. The society model exercise has revealed a number of potential tensions that could act as constraints, and therefore as potential challenges for further progress. This is important, as we do not expect our citizens to spontaneously use the system as it is offered to them. We must have a better idea about the existing barriers, not only between civilians and public administration, but also between all stakeholders and this particular application of social media. The following tensions were assumed:

- Users (both civilians and pa's) are unfamiliar with open data and with interpreting graphs
- The two groups of stakeholders are not familiar with this type of collaboration
- Civil servants are not allowed to speak openly on behalf of their professional organisation

- Employers are used to speak openly about local government, but not about their deals
- Private deals between employers and local government are important because they are private, they lose importance when they become shared
- Not all solutions work in the same way when they become open

We can say that the type of ideas and solutions that lend itself for open discussion on a social medium, or open to some extent, are of a different nature than idea and solutions discussed so far: they are more general, in the meaning of good for all, and more for the longer term, in the meaning of not immediately useful. This requires a different way of thinking and collaboration, and this involves longer-term, and group-oriented perspectives. At least, that is what we suppose here. The idea is that the more open stakeholders stand towards such types of ideas and solution, the more positive stance they will hold with regard to open data, co-creation and social media. We will investigate their evaluation and stances as they evolve with the appraisal framework (White, 2011; Martin & White, 2005). Within this framework, we expect positive affect (feelings), greater social acceptability and greater appreciation of transparency as it is given shape in open discussions between public administration and employers, in the upcoming period (M13-19).

Tentative timing of activities:

- 1) February 2016: workshop on collaboration in social media
- 2) March 2016: workshop on working with the tool, negotiation of first scenario
- 3) April-June 2016: scenario-based co-creation activities
- 4) April-June 2016: further selection and development of useful open data, recruitment of further participation
- 5) Second half of 2016: pilots with scenarios and several working groups
- 6) December/January 2016-17: Evaluation and D5.2 production.

REFERENCES

- Allwood, J., Traum, D. & Jokinen, K. 2000. Cooperation, dialogue and ethics. *International Journal of Human-Computer Studies*, 53, 871-914.
- Andriessen, J., Baker, M. & Pardijs, M. (2013). Getting on and getting along: tensions in the development of collaborations. In M. Baker, J. Andriessen & S. Järvelä (Eds.) *Affective Learning Together: Social and emotional dimensions of collaborative learning* (pp. 205-230). London: Routledge.
- Asterhan, C. S. C., & Schwarz, B. B. (2010). Online Moderation of synchronous e-argumentation. *Computer-Supported Collaborative Learning*, 5, 259-282.
- Baker, M.J. (2015). Collaboration in Collaborative Learning. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*, 16(3), 451-473.
<https://benjamins.com/?embed=1#catalog/journals/is.16.3/toc>
- Bereiter, C., & Scardamalia, M. (2003). Learning to work creatively with knowledge. In: E. De Corte, L. Verschaffel, N. Entwistle, & J. Van Merriënboer (Eds.) *Unravelling basic components and dimensions of powerful learning environments*. EARLI Advances in Learning and Instruction Series, pp. 55-68.
- Davis, F., Bagozzi, R., & Warshaw, R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, 35, 982-1003.
- Détienne, F., Baker, M. & Burkhardt, J.-M. (2012). Quality of collaboration in design meetings: methodological reflexions. *CoDesign: International Journal of CoCreation in Design*, 8(4), 247-261.
- Détienne, F., Baker, M., Fréard, D., Barcellini, F., Denis, A. & Quignard, M. (2016). The Descent of Pluto: interactive dynamics, specialisation and reciprocity of roles in a Wikipedia debate. *International Journal of Human-Computer Studies*, 86, 11-31.
- De Villiers, M. R., & Pretorius, M. C. (2013). Evaluation of a Collaborative Learning Environment on a Facebook Forum. *The electronic Journal of Information System Evaluation*, 16, 1, 56-70.
- Dreier, O. (1999). Personal trajectories of participation across contexts of social practice. *Outlines: Critical Social Studies* 1:5-32.
- Fréard, D., Détienne, F., Baker, M., Quignard, M., Barcellini, F. & Denis, A. (2012) Visualising zones of collaboration in online collective activity: a case study in Wikipedia. ECCE 2012, *European Conference of Cognitive Ergonomics*, Edinburgh, UK, August 28-31. <http://hal.inria.fr/hal-00766463/>
- Hew, K.F. (2011). Students' and teachers' use of Facebook. *Computers in Human Behavior* 27, pp. 662-676.
- Kraut, R.E. & Resnick, P. (2011). *Building successful online communities: evidence-based social design*. Cambridge: MIT Press.
- Lave, J. & Wenger, É. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.
- Ludvigsen, S., Rasmussen, I., Ingeborg, K., Moen, A. & Middleton, D. (2011). Intersecting trajectories of participation: temporality and learning. In S. Ludvigsen, A. Lund and R. Säljö (Eds). *Learning Across Sites: New tools, infrastructures and practices*, pp. 105-121. Routledge: London.
- Klaudia Thellmann, K., Galkin, M., Orlandi, F., Auer, S. (2015). LinkDaViz – Automatic Binding of Linked Data to Visualizations In Proc. of "The Semantic Web - ISWC 2015, Volume 9366 of the series Lecture Notes in Computer Science pp 147-162
- Maloney-Krichmar, D. and Preece, J. 2005. A multilevel analysis of sociability, usability, and community dynamics in an online health community. *ACM Trans. Comput.-Hum. Interact.* 12, 2 (June 2005), 201-232.

- Martin, J. R. R., & White, P. R. R. (2005). *The language of evaluation : appraisal in English*. New York: Palgrave, Macmillan.
- Mazman, S.G. and Usluel, Y.K. (2010). Modeling educational usage of Facebook. *Computers & Education*, Vol. 55, pp. 444-453.
- Meijer, A., Grimmelikhuijsen, S., & Brandsma, G. J. (2012). Communities of Public Service Support: Citizens engage in social learning in peer-to-peer networks. *Government Information Quarterly*, Vol. 29, Nr. 1, pp. 21–29.
- Miller, D.L. (2003). The Stages of Group Development: a Retrospective Study of Dynamic Team Processes. *Canadian Journal of Administrative Sciences*, 20(2), 121-134.
- Paavola, S., Lipponen, L., & Hakkarainen, K. (2004). Models of Innovative Knowledge Communities and Three Metaphors of Learning. *Review of Educational Research* 74(4), 557-576.
- Stanoevska-Slabeva & Schmidt, 2001 System Sciences. Proceedings of the 34th Annual Hawaii International Conference.
- Venkatesh V., Morris M.G., Davis G.B, & Davis F.D. (2003), User acceptance of information technology: toward a unified view, *MIS Quart.*, 27 (3), 425–478
- Wenger, E. (1998). *Communities of practice. Learning, meaning, and identity*. Cambridge: Cambridge University Press.
- Wenger, E., McDermott, R., & Snyder, W. M. (2002). *Cultivating Communities of Practice: a guide to managing knowledge*. Boston, Massachusetts: Harvard Business School Press.
- White, P.R.R. (2011). Appraisal. In Jan Zienkowski, Jan-Ola Östman and Jef Verschueren (Eds.), *Discursive Pragmatics*, pp. 14-36. Amsterdam: John Benjamins.