LIVE+GOV

Reality Sensing, Mining and Augmentation
for Mobile Citizen–Government Dialogue

FP7-288815

D5.3

Results of first trial and revised requirements

Dissemination level:	Public
Contractual date of delivery:	Month 25, February, 2014
Actual date of delivery:	Month 25, February, 2014
Work package:	WP5 eParticipation Use Cases
Task:	T5.4 Field Trials
Type:	Prototype
Approval Status:	PMB Final Draft
Version:	1.2
Number of pages:	142
Filename:	20140227_D5.3_Results_of_first_trial_and_revised_requirements_v1.2.docx

Abstract

In D5.3 we evaluate the customised prototypes as described in D5.2. For the evaluation, each Use Case executed field trials: Mobility, Urban Maintenance, and Urban Planning. In these field settings, our applications have been actively used by our target audiences. In this deliverable, we present the overall evaluation strategy to measure the usability of our prototypes, check our requirements as set out in D5.1 in field-settings, and report on the impact of our field trials. Overall, we present specific feedback for WP1-WP4 and describe the revisions of our functional requirements in order to prepare for our larger scale trials based on the second cycle of developments. With the evaluations, the Live+Gov consortium has gathered a tremendous collection of information for the next Live+Gov steps and is able to continue fortifying engagement and making impact.

Copyright

© Copyright 2013 Live+Gov Consortium consisting of:

1. Universität Koblenz-Landau
2. Centre for Research and Technology Hellas
3. Yucat BV
4. Mattersoft OY
5. Fundacion BiscayTIK
6. EuroSoc GmbH

History

Version	Date	Reason	Revised by
0.1	8-10-2013	Document outline set-up.	Pieter Minnigh
0.2	17-12-2013	Alpha version	Pieter Minnigh
0.3	7-1-2014	Internal comments and feedback incorporated	Pieter Minnigh Matthias Thimm
0.4	19-12-2013	Input WP 1 and WP3 incorporated	Pieter Minnigh Heinrich Hartmann Spiros Nikolopoulos
0.5	23-1-2014	Initial Urban Maintenance input incorporation, content writing	Pieter Minnigh Marjolein Terpstra
0.6	30-1-2014	Mobility, Urban Planning input incorporated	Pieter Minnigh Laura Niittylä Maite de Arana Agiretxe Jon Casasempere
0.7	6-2-2014	Input Mobility, Urban Maintenance, Urban Planning incorporated. Restructuring work.	Pieter Minnigh Laura Niittylä Marjolein Terpstra Maite de Arana Agiretxe Jon Casasempere
0.8	11-2-2014	Rework and added content on all sections. Input WP4 incorporated.	Pieter Minnigh Laura Niittylä Marjolein Terpstra Maite de Arana Agiretxe Jon Casasempere Frank Thiele
0.9	26-2-2014	Input Mobility, Urban Maintenance, Urban Planning incorporated. WP4 input revisions. Complete overhaul of the deliverable.	Pieter Minnigh Laura Niittylä Frank Thiele Maite de Arana Agiretxe Jon Casasempere Marjolein Terpstra
1.0	26-2-2014	Overall revisions, corrections, finalising	Pieter Minnigh Maite de Arana Agiretxe Frank Thiele
1.1	27-2-2014	Final fine-tuning	Pieter Minnigh Frank Thiele Marjolein Terpstra Jens Steensma
1.2	28-2-2014	Final editing	Matthias Thimm

Author list

Organisation	Name	Contact Information
YCT	P.A. Minnigh	**p.a.minnigh@yucat.com**
MTS	Laura Niittylä	**laura.niittyla@mattersoft.fi**
BIZ	Maite de Arana Agiretxa	**komunikazioa.tekniko1@biscaytik.eu**
YCT	Marjolein Terpstra	**m.terpstra@yucat.com**
BIZ	Jon Casasempere	**jon.casasempere@biscaytik.eu**
UKob	Heinrich Hartmann	**hartmann@uni-koblenz.de**
CERTH	Spiros Nikolopoulos	**nikolopo@iti.gr**
YCT	Frank Thiele	**f.thiele@yucat.com**
YCT	Jens Steensma	**j.m.steensma@yucat.com**

Executive Summary

This document *Results of first trial and revised requirements* is the report of the Use Case trial evaluations. For each of the three Use Cases, customised prototypes have been developed that have been used in the first field trials. In this document we present the overall evaluation strategy, the evaluations for each specific Use Case trial and the overall output of the trials in form of feedback for WP1-WP4 and revised functional requirements.

The evaluation strategy consists of measurement of user experiences in field-settings. We have divided the analysis in different parts for this evaluation:

- Usability aspects
- Requirements evaluations
- Feedback for WP1-WP4
- Impact of the trials

For usability, WP5 set out different acknowledged usability aspects, based on common interpretations of usability. These aspects are the following: Learnability, Utility, Memorability, Satisfaction, Efficiency, Feedback & Errors, and Reliability. In total, these aspects give a good impression on the overall user-experience of our applications. Furthermore, each Use Case presents the results of the evaluation of their functional requirements. Based on these evaluations, we report on revisions in these requirements and the current state of the implementations. WP1-WP4 have been consulted for specific information needs out of the field trials. The Use Cases conducted their trials taking these information needs into account and will make recommendations and provide feedback to the WPs. Finally, each Use Case reports on the impact and success of their field trial. User and usage numbers, media coverage and overall lessons learned for the next trials are presented.

The trial evaluations provide the consortium with diverse results. Conceptually, each Use Case can report successful results: our applications were met with very positive reactions by both sides of our targeted users (citizens and government). Individual improvements on usability aspects are assessed, but generally speaking these improvements are of minor impact. On the other hand, in both the Mobility and the Urban Planning Use Case, the dependency on technical stability and accuracy showed itself crucial. The field-settings were definitely more diverse and more demanding than was expected upfront. Both on the technical verification and the field-trial preparations, we will need to be more concise for the second field trials. Even though the more technical issues are challenging, the actual enthusiasm and strengthened stakeholder engagement dominate overall. For the Use Cases we are able to present impact in large usage numbers, fortified stakeholder commitment and further embedding in the organisations. Our stakeholders are confident in the Live+Gov approach and participatory applications. Interesting to see is that our solutions fuel the transition towards more citizens-centric society and the needed organisational discussions about the role of the government.

With the results of the first field trials, we report on our first end-user check within the project. With these evaluations, the consortium is provided with input for further technical development, shaped and fine-tuned functional requirements, insights on executing field-trials, and embedding of our solutions in organisations. With the evaluations, the Live+Gov

consortium has gathered a tremendous collection of information for the next Live+Gov steps and is able to continue fortifying engagement and making impact.

Abbreviations and Acronyms

AR	Augmented Reality
BIZ	Live+Gov Consortium partner Fundacion Biscaytik
CERTH	Live+Gov Consortium partner Centre for Research and Technology Hellas
DoW	Description of Work
Dx.x	Project Deliverable x.x
EuSoc	Live+Gov Consortium partner Eurosoc GmbH
GPS	Global Position System
HAR	Human Activity Recognition
HSL	Helsinki Region Transport
JMU	*Jij Maakt Utrecht*
L+G-FR.xx	Live+Gov Functional requirement xx
MTS	Live+Gov Consortium partner Mattersoft OY
Mx	Project Month x
RFID	Radio Frequency Identification
SaaS	Software as a Service
StUF	Standaard Uitwisselings Formaat (Dutch for Standardised Sharing Format)
Tx.x	Project Task x.x
UKob	Live+Gov Consortium partner University of Koblenz-Landau
UML	Unified Modeling Language
WP	Work Package
YCT	Live+Gov Consortium partner Yucat BV

Table of Contents

List of Figures

List of Tables

1 Introduction

This document *Results of first trial and revised requirements* is the public report D5.3 as described in the Description of Work (DoW). This document contains the evaluations of the first field trials of the three Use Cases: Mobility, Urban Maintenance and Urban Planning. In D5.1 – *"Detailed Use Case Description"* [2]each Use Case defined their specific requirements. These requirements were used to facilitate the development of the functionalities for WP1-WP4. With the use of the Live+Gov Toolkit components, the customised Use Case prototypes have been delivered in D5.2 [1]. D5.3 encompasses the evaluations of the first field trials with these prototypes.

In Figure 1 the planned general cyclic overview can be seen, and the phase of current D5.3 (and direct-related T5.4):
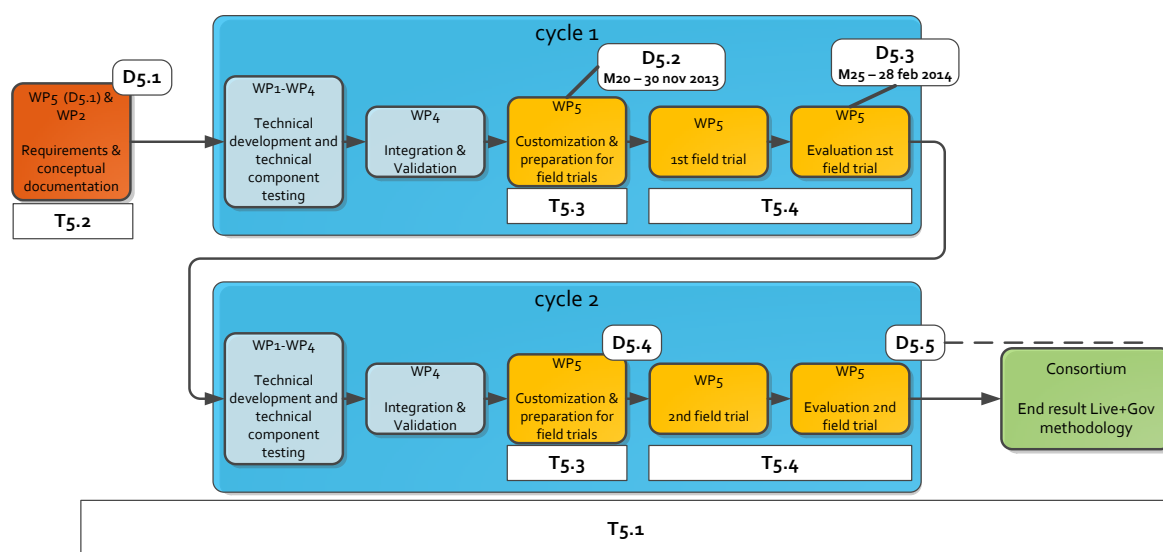
Figure 1: D5.3 & T5.4 in Live+Gov

1.1 What to expect in D5.3?

The DoW description for D5.3 (and the linked T5.4) sets the scope of this document and can be summarised as follows:

- The trials that will be conducted using the outcome of D5.2 will be evaluated in D5.3.
    - Technical results will be feed backed to WP1 to 4.
    - Overall results will also be used in the Dissemination Work package.
    - In-between results, that will be collected after the first trial run, will be used to improve the second run.
- The deliverable will report on the evaluation results and provide the revised requirements for the eGovernance augmented reality application.
- This task will perform the instantiation of the Use Cases, i.e. execute and evaluate actual field trials, based on the description set out in T5.2 and the customized prototypes of T5.3. During the field trials a close relation will be maintained with the relevant technical WP's and the methodology development in T5.1, in order to give

constant feedback that enables these Work Packages to further improve their models and the methodology to mature.

- The actual field trials will be done in the two phases (Phase III: Evaluation 1 and Requirements Check and Phase V: Evaluation II):
    - phase one will be performed with an initial prototype on selected subjects,
    - the second phase will have the full system and will be made available to a large audience.
- During field trials, both parties involved – citizens and governments – must be actively involved. This will be accomplished by using existing communities (for example the BuitenBeter community in the Urban Maintenance Use Case) and engaging (local) government that benefit from the results of the trial.
- The operation of the prototypes is accompanied by corresponding evaluations. The evaluations will be technical, organizational, and user-centric. In the definition phase, objective criteria will be set out that will be used during evaluation.

In D5.3 – *"Results of first trial and revised requirements"* we deliver in accordance with this description of the DoW, the following:

- The description of the evaluation strategy
- The evaluations of each individual Use Case Field Trial
    - Usability evaluation (WP5)
    - Functional requirements evaluation (WP5)
    - Specific recommendations and results for WP1-WP4
    - Impact and success of the trials (which is also input for WP6)
- Summary evaluation results of the Functional Requirements

In the next section, the structure of the deliverable that reflects the topics described is presented.

1.2   Structure of D5.3

This document is structured as follows: In the current section 1, the scope and the structure of this deliverable are described. In section 2 the evaluation strategy for the first field trials is defined by specifying the overall goals and the evaluation approach for each individual trial. This strategy brings results and insights forth for WP1-WP5. Specific attention has been given to the individual WP information needs. These needs are also described in section 2 and have been derived from consulting each WP-leader about his vision on the methodology that seems most productive for delivering the desired information. This shall provide the project with evaluation insights on the developed technologies from a functional/user perspective.

Sections 3, 4, 5 and 6 are the specific Use Case trial evaluations:

- Section 3 – Mobility Use Case
- Section 4 – Urban Maintenance: *Jij Maakt Utrecht*
- Section 5 – Urban Maintenance: Issue reporting and reality mining
- Section 6 – Urban Planning Gordexola

In each of these Use Case sections the trial is described, a summary of the customised prototypes that have been used during the field trials is given (as delivered in D5.2 [1]), and the tailored Use Case evaluation process is set. After these introductory descriptions, the actual field trial results and evaluations are presented. For each Use Case evaluation the usability of the application is studied following the WP5 strategy to specifically test Learnability, Utility, Memorability, Satisfaction, Efficiency, Feedback & Errors, and Reliability. The evaluation of usability aspects is followed by evaluating the functional requirements for the first trial and giving recommendations and revisions for these requirements. Based on the field trials, recommendations are also made for the other Live+Gov WPs. Finally, the impact of the field trials is analysed, based on user and usage numbers, media coverage and other aspects of impact.

After the individual Use Case sections, section 7 combines and summarizes the results of the individual cases. Here, the summary of the evaluation results of the Functional Requirements tested in the first field trials is presented.

2 Evaluation strategy

In this section, we describe the overall evaluation strategy for the first field trials, which consists of usability evaluations of our applications, user experience evaluations of the implemented functional requirements, specific results for WP1-WP4, and the success and impact of our first field trials.

2.1 User-experience evaluation for each Use Case

In our Use Case trial evaluations we will primarily focus on the aspect of user-experience regarding the customised prototypes described in D5.2[1]. These prototypes are the first 'end results' of the Live+Gov developments and by assessing the quality of experiences gained by our targeted audiences we practically aim to check on the validity of the requirements specified in D5.1 [2], as well as the implementations developed to address these requirements. User experience, however, is not trivial to grab without well-defined aspects to measure. WP5 set up the evaluation strategy for the field trials and deconstructed usability into seven usability aspects. In 2.1.1 the aspects, which are acknowledged to be key aspects for user experience measurement, are described. It is important to state that in WP4 the procedures for the technical verification of the software are specified (see D4.4 [3]). In WP5 we validate that what has been built meets the end-users needs. Therefore, we evaluate our requirements as set out in D5.1.[2]

2.1.1 Usability

One of the main concepts for the evaluation of user experiences is the software usability. The definitive formulation of usability is hard to find agreement on, however large overlap can be found on the different usability aspects. WP5 provided the Use Cases with survey templates and usability guides based on 9 different usability questionnaire survey standards (see Table 1) that we've analysed.

Acronym	Instrument	Author, year, organisation
QUIS	Questionnaire for User Interface Satisfaction	Chin et al, 1988, Maryland
PUEU	Perceived Usefulness and Ease of Use	Davis, 1989, IBM
NAU	Nielsen's Attributes of Usability	Nielsen, 1993, Bellcore
NHE	Nielsen's Heuristic Evaluation	Nielsen, 1993, Bellcore
CSUQ	Computer System Usability Questionnaire	Lewis, 1995, IBM
ASQ	After Scenario Questionnaire	Lewis, 1995, IBM
PHUE	Practical Heuristics for Usability Evaluation	Perlman, 1997, OSU
PUTQ	Purdue Usability Testing Questionnaire	Lin et al, 1997, Purdue
USE	USE Questionnaire	Lund, 2001, Sapient

Table 1: Usability instruments [8]

From these standards, we distinguished seven different usability aspects as can be seen in Table 2. By applying this standardised interpretation of usability, we make sure that each Use Case gets a broad understanding about the user experience of the prototypes.

Aspect	Description
Learnability	How easy is it for users to accomplish basic tasks the first time they encounter the user-interface?
Utility	Does it do what users need?
Memorability	When users return to the user-interface after a period of not using it, how easily can they re-establish proficiency?
Satisfaction	How pleasant is it to use the system in general?
Efficiency	Once users have learned how to operate the user-interface, how quickly can they perform tasks?
Feedback & Errors	How many errors do users make, how severe are these errors, and how easily can they recover from the errors?
Reliability	Does the software run consistently without crashing?

Table 2: Usability aspects

Along with these aspects, the prototypes and the foreseen requirements will be evaluated by each Use Case owner for their respective Use Case. What measurement method fits the purpose best is to be assessed by each Use Case owner and depends on the trial set-up. In general, the way each Use Case will gather the information is a variation and combination of survey questionnaires, interviews, analysis of data from of systems 'databases or user-activity logs. The methods applied for the Use Case evaluation are described in the respective Use Case sections.

For our evaluation we assess the usefulness of our results for the individual WPs as crucial for further development. Each Use Case evaluation will lead to concrete recommendations, information and improvements for the different WPs. To be sure that specific information needs are not overlooked by the evaluation process, we included the input of technical WPs (WP1-4) in the strategy for the evaluation in order to obtain feedback that is as productive as possible.

Important to note here is that research on usability samples shows, that a large sample for usability tests is not necessary. Moreover, a sample of more than 5 test users gains little to no extra insights (see Figure 2). However, if a system is used by completely different types of users and the functionality of the system is also different, a larger group is necessary: the system should be approached as if it were a combination of different individual systems. [9] For example, 10 test users are necessary if the system has two different types of users.
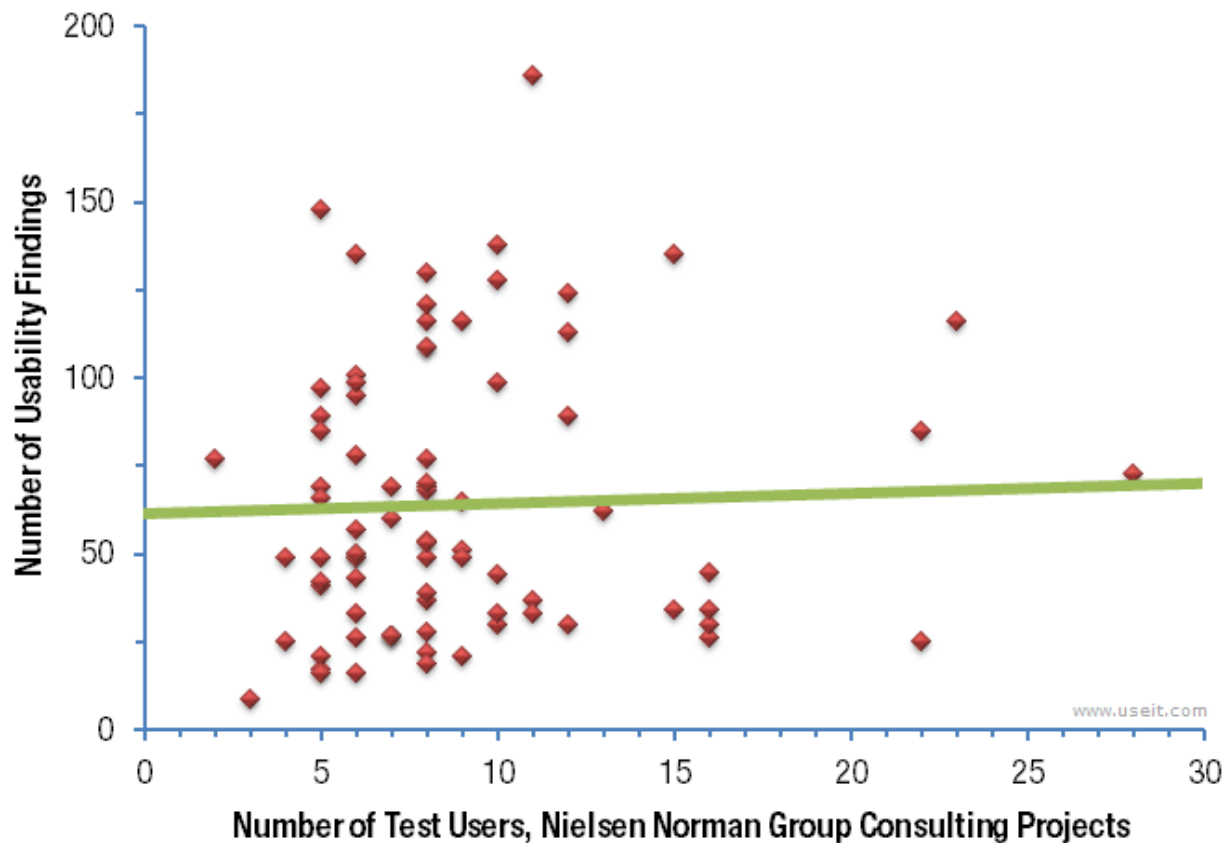
Figure 2: Summary of usability consulting projects research: findings vs test users [9]

2.1.2   Requirements

The key aspect of our cyclic Live+Gov evaluation strategy is to check the developments against the requirements. With the trial evaluations, we essentially evaluate the requirements from a functional user-perspective. The evaluation can lead to different outcomes (here described in a negative fashion for clarity) regarding the requirements:

- End users aren't happy with the way the requirement is *implemented*. This can for example relate to issues of:
    - (technical) development of components that actually don't fit neatly in the field situations
    - The customisation of the prototype itself, including GUI design, is not fully appropriate for our targeted users.
- Requirements can be *defined* not precise enough, which leads to 'wrong' development.
- The *assessment of the aims* of the requirements wasn't correct. This could be that the definition of the underlying 'problem' wasn't clear, but can also relate to wrong assumptions consequently leading to wrong requirements. The definition itself is not wrong, the aim of the requirement can prove itself to be incorrect.
- The requirements are not necessarily *complete*. Important requirements are simply *missing*, which leads to poor user experience due to incomplete solutions.

The evaluations of the first field trials are based on the user experiences of our Live+Gov prototypes with full functionality. However, we are in the first cycle of developments and the second cycle is more than just shaping and shaving on our development: some important

innovations are planned for the upcoming period and cannot be the subject of our current evaluations.

2.2 Specific WP evaluation information needs

To make sure that the field trials are used optimally to gather insights for future steps in Live+Gov, WP5 set up specific templates and communicated intensively with WP1-WP4 so as to assess the information needs. The perspective of the field trial evaluations is for the most part functional, however, functionalities are offered by the technical components. For the trials, these components have been combined into fully customised trial prototypes (D5.2, [1]) and the combination of these functionalities they offer will therefore be checked. The components have been developed based on the Use Case requirements as specified in D5.1 [2]. In the first field trials, we are able to do the check on the requirements themselves, the offered functionality and the implementations. The trials are a good source of information for assessing the behaviour of prototypes and components in dynamic real-life settings. It is important to note that, in the context of the field trials, our aim is to evaluate the developed technologies from a functional/user perspective and not using objective performance metrics, i.e. for example the performance for rendering 3D models, or the accuracy of service line detection.

It is important to stress once more that for the first field trial, not all the assessed functionalities have been fully developed yet or are in their final form. Some components are mainly targeted for a specific Use Case; based on results of the first trials we are able to shape and extend these components for inclusion in other Use Cases. One clear example is the Augmented Reality component and the functionality of triggering meta-tags using image recognition. For the first cycle, these features are included in the Urban Planning trial, but for the second cycle we foresee inclusion in the Urban Maintenance Use Case as well.

In the next sections, we describe the information needs for each of the technical WPs (WP1-4) with respect to the evaluation process.

2.2.1 WP1 – Reality Sensing and Mining

The main aim for WP1 in the first field trial is to test the functionality of the delivered components under field conditions. These components are the following:

1. Sensor Data Capturing Service (S1)
    a. Mobile Sensor Collection Component (C14)
    b. Sensor Data Storage Service (C8)
2. Reality Mining Service (S2)
    a. Server Side Mining Service (C9)
        i. Service Line Detection Component
        ii. Traffic Jam Detection Component
    b. Mobile Sensor Mining Component (C15)
        i. Human Activity Recognition Component
3. Initial Route Analysis Tool

In the evaluation of the use case we want to judge if the delivered components meet the functional requirements on a user level. This stands in contrast to a technical evaluation of the classifiers (e.g. precision recall measurements, confusion metrics) that will be performed in D1.2 (*Sensor Data App with Mining Functionality*, forthcoming) and component testing

which is addressed in D4.4 [3] and D4.3 (*Report on the Phase 2 Integrated Live+Gov toolset and AIV Test Report*, forthcoming).

2.2.1.1 Evaluation of Sensor Data Capturing Service

The delivered component shall be able to record GPS tracks and raw sensor data from the accelerometer sensor and transfer them to the server. This functionality has been tested and reported in D1.1. [4] We have an inspection tool in place, which gives a concise overview about the recorded routes and the quality of the recordings.

We expect to be able to store all routes that are travelled during the trial, and uploaded by the user. To evaluate this requirement, we will collect error messages of failed uploads on the server and on the mobile device. Moreover, a plausibility calculation will be performed comparing the number of recorded routes to the number of participants and duration of the trial.

2.2.1.2 Evaluation of Reality Mining Service

Progress on the Reality Mining Service by WP1 for the first field trials includes the initial Server Side Mining Service and the Human Activity Recognition (HAR) component. The progress made is in accordance with the DoW descriptions, but it should be noted that for the first trials this progress is centralised on the Mobility Use Case. Next steps in knowledge discovery by data mining, is planned for the upcoming period. This gives us the advantage to utilize the gained experience and results from a single case in order to realise broader inclusion in Live+Gov.

Server Side Mining Service

For the first trials, we have the Service line Detection Component and the Traffic Jam Detection component in place. These support the functional requirements as assessed in D5.1:

- Be able to sense the context of users (e.g. location, means of transportation, activities etc.)
- Receive data from the application containing travel information and means of travel.
- Detect traffic jams

The results of the service line detection service are directly shown to the user. First feedback about the quality of the detection will be collected directly during the trial. For this purpose a feedback mechanism is included into the integrated application. When the user clicks on the detected service line, a list is shown, where the user can select the correct service line that he is currently using.

The evaluation of the Traffic Jam Detection Component is done based on user experience. The trial users are asked how well the reported jams have reflected the traffic situation. As a secondary evaluation method, manual checks are done. A log of detected jams is created and from the logs a manual check of situations where jams are known to have happened is done. Time and location of a known jam is compared to the logs to detect if alerts are given of these situations.

Evaluation of Human Activity Recognition (HAR) Component

A first basic evaluation of the HAR component would be to present the recognized activities to the user and ask them for their experience about the accuracy. The integrated mobility-

application shows an iconic representation of the recognized activity in the front end. However, there is a principal difficulty with this approach: namely, the user needs to hold the device in his hand in order to view the icon, thereby disturbing the HAR classifier, which is trained to recognize activities when worn in a pocket. During the trial, we will use an internal data-inspection tool for accuracy checks; this is a technical evaluation to be able to get insights on accuracy in a field setting. The inspection tool will show the classified activity alongside with the recorded routes.

2.2.1.3 Evaluation of the Initial Route Analysis

The quality and the usefulness of the route analysis will be evaluated in interview-sessions with the local authorities in Finland (HSL). Here, it is important to gather insights in the needs on specific information and presentations. In the first field trial, this will include analysis of journeys in a given time frame and journeys crossing a point near 'x'. The internal inspection tool will become helpful to have an initial idea for data-presentations in the Mobility Use Case.

2.2.2 WP2 – Policy Modelling and Visualization for Mobile-aware Dialogue

For WP2, the Field Trials can provide invaluable insights for constructing the Training Package as one of the end results of Live+Gov. This handbook is meant to give organisational bodies tools and best practises about how to start and implement participation scenarios. The overall evaluations and insights of the Use Case trials will generate information on current practises and challenges on a political-organisational level. As described in D2.3, society is in the transitional process towards a more we-government approach. In participation-societies, informal decision making procedures play a way more important role than in more traditional models. The process towards the participation society depends on the success of implementing visualisation techniques to structure informal decision making processes [12].

From the field trial evaluations, we will get information on the difficulties and challenges experienced in bringing structure on informal decision making processes as well as the organisational adaption of eParticipation solutions as our use cases. Important feedback for WP2 from the field trials can be linked to the following two aspects:

- Organisational insights and lessons on adapting eParticipation scenarios in (local) governments
- Insights on visualisations in order to structure informal processes of decision making. The role of these becomes more and more important in the transition towards the participatory organisational model.

2.2.3 WP3 – Mobile Augmented Reality

The role of WP3 in the general context of Live+Gov is to provide the necessary means for engaging citizens in a two-way dialogue with their government. There have been three main lines of action for reaching this goal:

a) Offering the citizen an enhanced view of the reality by augmenting the environment around him with government-related information,
b) Allowing the mobile user to benefit from services like image recognition and personalized content delivery, and

c) Processing massive user contributions and producing aggregated views and comprehensive summaries for following-up on the governmental issues.

Based on the above, our focus during the first field trials of the project is to assess the level of engagement achieved through the use of mobile augmented reality, as well as evaluate how much value the users will find in using image recognition and how much value the users will find in the personalization services.

Towards this end, we utilize two main instruments:

- Questionnaires that capture the user opinions explicitly through a set of questions,
- Mobile app usage logs that are able to implicitly derive users' preferences, likes and dislikes.

2.2.3.1 Level of engagement achieved through Augmented Reality

Location-based augmented reality is typically used as an alternative to map-based interfaces, where instead of pin-pointing the information on a map, it is overlaid on the camera view of the mobile phone. When this information relates to infrastructure plans, or designs for future constructions the combination of mobile augmented reality with 3D rendering technologies are able to offer the citizen a much more realistic view of the governmental plans. This is exactly the scenario that has been implemented in the prototype mobile application of the first field trial for Urban Planning. The citizen is offered the possibility to view future governmental plans either through a map-view, a list-view or an augmented reality view that incorporates the 3D models of future constructions. Our interest, from the perspective of WP3, is to assess the level of citizens' engagement with the AR view compared to map- and list-based views, as well as the level of citizens' satisfaction.

2.2.3.2 Usefulness of image-based recognition

In contrary to location based augmented reality, the core concept of image based augmented reality is not to augment the user's environment based on his location, but based on what he sees around him. Thus, this type of service is appropriate for objects that do not have a fixed GPS location but can trigger the augmentation process based on their visual form and appearance (e.g. logo, drawing, etc). This functionality has been incorporated in the prototype application of the first field trial of the Urban Planning use case. Our interest in this case is to assess the amount of usefulness that the citizens find in this functionality.

2.2.3.3 Usefulness of personalized content delivery

The employment of personalized services in Live+Gov has been motivated by the need to filter out the information that is irrelevant to the user in the situation at hand. For instance, in the scenario prescribed by the Mobility use case it is very likely for the citizen to get overwhelmed by traffic-related information about the entire city that is of little practical use for his indented journey. This problem becomes even more evident in the small display of mobile phones where information should be communicated wisely. For the purposes of the first field trial of the mobility use case, the personalization aspect has been incorporated at the level of notifying the citizen on whether the tram-line that he is currently using is moving towards a traffic jam. WP3 interest in this case is on assessing the acceptance of the personalized traffic services by the citizens commuting with public transportation means.

2.2.4   WP4 – System Integration

The main goal for WP4 is to provide a Software-as-a-Service (SaaS) infrastructure for mobile applications, which offers readily development of mobile governance solutions in a sustainable, scalable and low cost way.

In order to reach this goal, an architecture is designed (see D4.1 - Report on Live+Gov toolkit requirements and architecture) [11] that follows a distributed approach supported by a central SaaS Service Center, which offers 'out of the box' functionality that is required in every typical SaaS-based mobile governance solution. In the first trial services are offered for account management, access control and diagnostics. Services for monitoring and billing are foreseen in the design, but not used in the first trial round.

The following aspects will be accessed during the first trial round to further improve the architecture, SaaS infrastructure and services for the 2nd trials.

2.2.4.1   Readily development / Cost-effective

During the development stage developers can profit from the existing tools in the Live+Gov Toolkit and the central services offered by the SaaS Service Center. Using these pre-built components and services should speed-up the development process for new mobile governance solutions. Two main evaluation aspects are identified.

The first aspect is the *required time / cost*:

- How much time is required for the integration of new tools with the SaaS Service Center?
- How much time/effort is required for the integration of mobile governance solutions with the SaaS Service Center?
- How much time is saved by using existing tools and central services instead of implementing their own mechanisms?

Time spent on integration tasks is obtained from the partner's time spending figures. Please note that this is limited to the actual time spent on integration and customization using the Live+Gov Toolkit. The alternative option (without toolkit), is not piloted within the Live+Gov project. However, the time spent using the toolkit can be compared with estimations based on the experience from the SME's in previous similar projects.

The second aspect is the *SaaS requirements*:

- Is additional SaaS functionality required within the existing central services? Were all use-case solutions successful in accomplishing their SaaS requirements, or did they require custom additions for the existing SaaS services?
- Are additional services required in the Service Center? The developed use case solutions will be analysed to check if two or more solutions contain similar (duplicate) functionality that can be centralized in future releases.

These questions are more focused on the development and customization phase, the subject of this evaluation is therefore not the trial user group, but the developers. WP4 will analyse the final use-case specific solutions and revised requirements to determine if additional SaaS functionality and/or services are required.

2.2.4.2   Scalability / Availability

Especially for SaaS based solutions scalability and availability are very important aspects. SaaS solutions only can become profitable if they can be exploited on a larger scale. Two aspects are important when assessing scalability: multi-tenancy and performance.

Multi-tenancy means that the system should allow addition of new users and customers with only minor effort. D4.1 describes different levels of multi-tenancy, where the aim is set for configurable and multi-tenant solutions. This means that a single instance of the solution is able to serve multiple customers, with each customer's data kept separately. Customers individual needs are still met by providing configuration options though account management and access control. [11] During the first trial run three separate mobile governance solutions are developed based on the Live+Gov Toolkit. Each use-case trial defines only one customer, therefore multi-tenancy will be applied on a higher level. During the trials the three different solutions will be served from a single instance of the SaaS Service Center. Furthermore this aspect will be verified during the integration- and the system-testing phase will verify that the individual SaaS solutions are capable of serving multiple customers on the same instance. The results of these tests will be reported in D4.3.

Performance is related to responsiveness and availability. Adding more customers and users will result in higher system load. During the trials the SaaS solution will be put to the test with user groups varying from 10 to 5000 people. The performance and availability will be accessed by 1) using objective measurement tools like Google analytics and 2) subjective user questionnaires to capture the actual user experience.

2.3   Impact and success of the Field Trials

One topic that hasn't been addressed yet is the impact and success of the trial itself. What is meant here is the scale, the execution and the user participation in the field trials. This type of information can benefit Live+Gov in two ways:

- With lessons learned of the first field trials, we will be able to have our second field trials set-up and executed more effectively
- On a higher level, the field trials resemble the broader participation difficulties of contribution to participation. What we experience on a small scale on involving stakeholders is on a bigger scale even more crucial.[1]

The impact and success of the trial do not influence the evaluation results, but they will have consequences for the credibility of the results. Apart from this, the success of the field trials will reflect in dissemination activities as well (WP6).

For the impact and success, we will report on user and usage numbers. Also, media attention will be described. If applicable, embedding of the use case in organisations is also considered as impact and consequently will be presented. Apart from these aspects, each use case will report on insights and lessons learned from the set-up and execution of their field trial.

---

[1] See for more details and background on the problem of involved costs (time, money and effort spent by citizens and authorities) and the burden for participation in D2.1.[5]

3   Mobility Use Case

3.1   Mobility Field Trial

3.1.1   Field Trial set-up

The mobility trial started on 19th of December 2013 and took place until January 31st 2014 in Helsinki region, Finland. In the trial volunteered trial users living in Helsinki region were provided a prototype of the mobility application. The volunteered users were asked to use the application during their daily travelling with public transport and share their impressions.

The application integration was done by CERTH and user recruiting and guidance by Mattersoft. From what was described in D5.2, some minor modifications to the application layout was done along the integration process but all required features and functionalities were kept in the application. Thus the setup of the trial can be seen to meet the pre-defined requirements.

The trial user recruitment started late October 2013 by Mattersoft. The primary channel for finding voluntary users was by a newsletter in HSL intranet, where it was visible to all HSL personnel. This method resulted in 10 volunteers to express their interest to participate by the given deadline of 29th November. As the aim was to gather at least 15 volunteers, not enough voluntary users from HSL were found with this method and the call was extended by two weeks. A number of external users were also recruited via public call and by contacting suitable users in the social media. These actions resulted in 15 volunteers in total by the start of the trial so the aim was fulfilled.

The original aim of having all trial users from within HSL was not met, but despite this, all volunteered test users were those in close contact either with HSL or Mattersoft personnel. The test user recruiting process turned out to be more challenging than expected and more focus on this needs to be addressed in the second trial.

As the trial started, all users were provided with a user manual in Finnish language. The manual provides step-by-step instructions on the application use. The application was delivered to the users via Google Play store, where the application was downloadable only on the Finnish market to prevent form external users downloading the application and distorting the trial results.

Support to the users was provided by Mattersoft during the launch and throughout the trial. Users reported some problems, asked several questions about the application use and the whole project. Answers were given and issue solving was done based on these contacts throughout the field trial.

3.1.2   Mobility prototype

For the implementation of the prototype the following requirements were defined for the first trial out of the complete use case requirements defined in D5.1:

The prototype shall

- Provide communicative tools for government and citizens (including reactions, alerts, messages (L+G.FR.05)
- Be able to capture images (L+G.FR.07)

- Be able to utilise 'citizens as sensors' for their environment by mining relevant data (e.g. location, context, pictures, reports) from their reality (L+G.FR.10)
- Be able to connect to different municipality administration systems (L+G.FR.13)
- Receive data from the application containing travel information and means of travel (L+G.FR.14)
- Detect traffic jams (L+G.FR.17)

The mobile prototype used in the trial has been described in detail in *'D5.2. Prototypes – demonstrators for first trials'* with step-by-step descriptions on all features and functionalities [1]. The prototype includes possibilities for the citizen to:

- record and submit data regarding their journey and provide additional details such as mean of travel and possible service line
- receive and view alerts and messages from authorities
- receive and view information on jams and exceptions
- send reports about issues in public transport

In the following table we summarize the Live+Gov components that have been employed in the 1st field trial of the Mobility Use Case:

#	Live+Gov component	Description
C1	Issue reporting Service	The issue reported is subsequently uploaded through the issue reporting service.
C2	Issue distribution service	Messages and alerts are entered via issue reporting back-end or web application and subsequently delivered to all users.
C8	Sensor Data Storage Server	Data collected from the sensors is stored in a flat file along with time-stamping information. Upon request of the mobile base application all data are uploaded to the Sensor Data Storage Server and removed from the mobile phone
C9	Server Side Mining Service (Service Line Detection Module)	Based on a series of GPS coordinates obtained from the citizen's mobile phone, identify the tram-line that user is currently travelling on.
C11	Issue reporting Client	The user is provided with a mobile interface to report an issue about his environment
C14	Mobile Sensor Collection Component	The raw data values are collected from the sensors of the mobile phone.
C15	Mobile Sensor Mining Component	The application is able to automatically recognize the user's current state of activity
	Google maps API	Tram lines and stops, as well as user routes and activity types are drawn on a map

	Traffic jam detection module (MTS)	Traffic jam detection module analyses actual tram location against schedules to detect jams
	Public Transport System (HSL)	Static information about the tram-lines and stops.

Table 3: Mobility Use Case: Live+Gov components

3.1.3 Evaluation process

In the evaluation of the use case we want to see if the delivered components meet the functional requirements on a user level, as they have been described in D5.1. The objective evaluation of the components and component performance will not be included in detail in this evaluation as these are addressed in D1.2, D4.4 and will be addressed in D4.3 [2][4][3]. In the evaluation we focus on the trial setup, user experiences, prototype usability and the trial impact and derive results for different work packages and also include the first impressions of the accuracy of provided information.

The evaluation of the trial is based on combined results from six different methods: Feedback from end users via questionnaires, personal contact with end users via e-mails, personal interview of HSL representatives, comparison to existing tools and techniques, data verification and Google Analytics statistics of the application usage.

In the middle of the trial the users were sent a first questionnaire, which aimed to provide insight on first impressions about the usability of the application and some information about user activity rates. After the end of the trial another questionnaire was sent, including more detailed and numerous questions regarding the trial and application. For the evaluation the questionnaires are the most important source for collecting end user experiences of the mobile application.

End user experiences are also collected through personal contact with the trial users. The main method for personal contact during and after the trial was by e-mails between the trial users and Mattersoft personnel.

Data analysis and internal reviewing is used to determine how the customization process succeeded and what level of user activity rates were reached throughout the trial. User activity is determined with the help of Google Analytics tool, which was implemented at the beginning of the trial to collect information about user behaviour and statistics. Data analysis also plays an important role in defining the quality aspects, such as the accuracy of HAR, service line detection and traffic jam detection.

After the trial, an evaluation meeting with HSL was also organized. In the meeting statistics of the trial was presented, setup of the trial, the mobile application and the included web interfaces were evaluated. As the first trial was defined to focus on the mobile application use, the web interface evaluation and impacts were done looking into the future based on what was delivered and how they are further planned to be developed. At this stage also the usefulness of the tools to be delivered based on current plans was evaluated.

The following Table 4 summarises all evaluation aspects for this use case. Each aspect includes specifying questions, used evaluation methods and the WP the evaluation of the specific aspect is delivering results to.

Prototype evaluation				
Aspect	Criteria	Specifications	Targeted WP	Methods used
Customization process success	Integration	How well were included components integrated?	WP4 & WP5	Internal review
	Application	Were all given requirements met?	WP5	Internal review
User activity	Quantity of data	Was enough data collected?	WP1 &WP5	Data analysis
	Interest	Did enough users continue regular use throughout the trial?	WP5	Data analysis, questionnaires, personal contact
	Coverage	Did enough users take part	WP1&WP5	Data analysis, questionnaires, personal contact
Usability	Learnability	How easy is it for users to accomplish basic tasks the first time they encounter the design?	WP5	questionnaires, personal contact
	Utility	Does the prototype meet the given requirements	WP5	Internal review
		Does it provide useful information?	WP5	questionnaires, personal contact,
		Is the personalized information useful?	WP3&WP5	questionnaires, data analysis
		Does it do what users need?	WP5	questionnaires, personal contact
	Memorability	When users return to the design after a period of not using it, how easily can they re-establish proficiency?	WP5	questionnaires, personal contact
	Satisfaction	How pleasant is it to use the system in general	WP5	questionnaires, personal contact
	Efficiency	Once users have learned the design, how quickly can they perform tasks?	WP5	questionnaires, personal contact
	Feedback & Errors	How many errors do users make, how severe are these errors, and how easily can they recover from the errors.	WP5 & WP4	questionnaires, personal contact, data analysis
		Feedback from the system to user	WP5	questionnaires, personal contact, data analysis
	Reliability	Does the software runs consistently without crashing	WP5 & WP4	questionnaires, personal contact, data analysis

Quality	Data	Is the data quality satisfying	WP1 & WP5	data analysis
	Traffic jam detection	Does traffic jam detection deliver accurate warnings	WP1	questionnaires, personal contact, data analysis
	HAR	Does human activity recognition deliver accurate information	WP1	questionnaires, personal contact, data analysis
	Service line detection	Does service line detection deliver accurate information	WP1	questionnaires, personal contact, data analysis
Impact	General	User numbers	WP5	Data analysis
		Usage numbers	WP5	Data analysis
		Media coverage	WP5	Internal review
	Authority	Does the application provide a useful tool for planning?	WP5	Interview, comparison to existing tools
	Citizen	Does this have an impact on citizens' mobility?	WP5	questionnaires, personal contact
	Interaction	Does the application provide a useful tool for interaction?	WP5	questionnaires, personal contact

Table 4: Evaluation aspects and methods in the first mobility trial

3.1.4 Usage of the application

Even though only the minimum anticipated number of users was recruited, the results of the trial have been positive in terms of application usage and amount of data transferred. The average activity rate of the users in general resulted in a positive surprise as higher than expected usage rates were gained on average even though some users have showed only minor activity. The following table shows the statistics of the usage during the trial based on Google Analytics.

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7
Active users	15	6	7	13	19	19	16
Sessions	173	118	109	358	343	376	307
Views	375	200	195	803	650	769	537
Views/Session	2,17	1,69	1,79	2,24	1,90	2,05	1,75
Average duration of a session	55sec	28sec	32sec	44sec	35sec	59sec	70sec
Crashes	20	0	2	3	6	23	10

Table 5: User activity during the trial

As can be seen, the number of active users has been growing during the trial, only exceptions being the weeks of Christmas and new year's holidays but the lack of use was expected on these weeks. Also during the final week activity has been moderately lower due to the end of the trial, taking place on Friday instead of Sunday. The user numbers

presented in the table are higher than the actual number of trial users due to having included also the test devices of Live+Gov project members (4 devices), that were used mainly in weeks 1, 5 and 6 when new updates were launched in the application.

When evaluating the user activity rates it must be taken into account that the participation was based on voluntary use. This also explains not all users providing answers to the questionnaires even though they have participated in the application use, out of the 15 trial users 10 users provided answers to the mid-questionnaire and nine users provided answers to the final questionnaire. As the used application at its current state still lacks several features that will increase usefulness and thus increase the interest to use, it would not have been realistic to expect any higher activity rate. Therefore, user wise the trial so far seems successful, number of active users reached an approvable number and the number of recorded journeys and usage time in total has exceeded the amount originally seen as the minimum for a successful trial.

3.2 User-experiences and customised prototype evaluation

3.2.1 Evaluation of Usability

In this section we analyse in more detail the user experiences on the mobile application. The results are summarised from the answers field trial users provided to the questionnaires and also based on the given comments throughout the trial. The citizen experiences on the trial focus on the usability of the prototype and the impact of the trial. In the figures of the tables presented in the sections below, the results are given as percentages of participants and the following scaling is used unless otherwise stated:

   0 = no opinion
   1 = strongly disagree
   2 = moderately disagree
   3 = moderately agree
   4 = strongly agree

3.2.1.1 Learnability

The trial users were asked to evaluate how easy the application is to learn to use and how intuitive is the navigation between the different functionalities is. The results show that users are satisfied with how easy the application was to learn to use, they are also pleased with the clarity and rationale of navigation.

	Question	Average	0	1	2	3	4
Learnability	I learned to use the application quickly	3.2	0 %	0 %	10 %	50 %	40 %
	Navigation is clear	3.3	0 %	0 %	0 %	70 %	30 %
	The application is easy and effortless to use	3.1	0 %	0 %	20 %	50 %	30 %
	Navigation is logical	3	0 %	0 %	20 %	60 %	20 %

Table 6: Mobility user opinions on learnability

Based on the numbers in table 5 we can state the learnability of the application receiving positive results. No severe points of improvement were found and generally the users were satisfied with the application. However, minor modifications and focus on the usability should still be included at on the second trial as more features will be included..

3.2.1.2 Utility

For the future development of the application, it is also important to evaluate how the users found the utility of the application, would it provide useful information and would it be something they would use in their daily travelling.

At its current stage, users mainly stated the application not being useful enough so that they would use it daily. Many commented that additional information should be provided for the application to be useful. Many of the requests however were on the type of information that is planned to be included in the second trial. These requests included e.g. providing more personal information and possibility to receive information before the journey starts. Also an application with personalized traffic information was agreed to be useful. Therefore these can be seen as positive results on the usefulness of the application in total.

Also comments on usefulness of features that are not planned to be included in the use case were given. Some users would have liked the application to show the real-time location of vehicles on a map or to provide suggestions with alternative routing on the application. These features are falling further off the project scope and are planned not to be included. However, the comments need to be addressed and evaluated in more detail to see if these features would be bringing value to the use case also in other terms than increased feeling of utility for the users.

	Question	Average	0	1	2	3	4
Utility	The application provides useful information	2.8	10 %	10 %	20 %	40 %	20 %
	The application provides a tool that is useful	2.6	10 %	0 %	40 %	40 %	10 %
	Personalized traffic information is useful	3.6	0 %	0 %	0 %	40 %	60 %
	The application provides useful information for public transport user	3.1	11 %	0 %	11 %	56 %	22 %
	The application provides a useful tool for public transport user	2.8	0 %	0 %	33 %	56 %	11 %
	The application provides a useful tool for interaction	3.1	11 %	0 %	0 %	78 %	11 %
	The application improves my chances of receiving useful information regarding my journeys	2.6	11 %	11 %	22 %	44 %	11 %

Table 7: Mobility user opinions on utility

In terms of utility, there are several important points of improvement detected based on the first trial results. However, these are all mainly the types of improvements that were already planned to be implemented in the second trial and the results were as expected. The most significant point to be taken into notice is to search the possibilities for the user to filter unwanted information to increase the feeling of usefulness of the application.

3.2.1.3   Memorability

When asking the users their opinion on the memorability of the application after a situation where there was a long period of not using the application, the general opinion was positive and the users stated the memorability being good. Only few users stated minor problems in the memorability. Therefore in terms of memorability the evaluation reached positive results with no major points of improvement for the second trial to be noted. However, this is an important factor also for the user satisfaction and should not be left without consideration also in the development of the second trial prototype.

	Question	Average	0	1	2	3	4
Memorability	After a period of not using the application, I find it likely to quickly to re-learn / remember how to use the application	3.4	0 %	0 %	22 %	11 %	67 %

Table 8: Mobility user opinions on memorability

3.2.1.4 Satisfaction

The design of the application was considered good by the users. It was found easy and satisfying to use in terms of layout and look and feel. However, it was stated that it would be more pleasant to use if the application started the recording automatically as the users reported often starting their journeys in a hurry and not remembering to push the '*start recording*'-button. Most users also agreed that the use of battery life was satisfying and did not affect the usability or satisfaction.

Major factor to cause dissatisfaction with the users was the number of alerts, which some found disturbing. Even though alerts were seen useful, it was stated that it would be useful to be able to determine whether the alerts and messages would be pushed to the user or if they would only be shown when user wanted to view them on the application. Also the manual starting of the recording was found difficult and some users stated hoping for an automatic starting of the journey detection in the application. When asking about the users' opinion on how satisfying the use of the application was, a few users regarded the design and look and feel of the application being good but due to constant crashes in the application they rated the satisfaction very low.

	Question	Average	0	1	2	3	4
Satisfaction	I liked the design of the application	3.1	0 %	0 %	10 %	70 %	20 %
	I'm satisfied with the battery usage of the application	2.5	10 %	0 %	50 %	40 %	0 %
	On a scale from 1-10 how satisfying the application use was?	6.4					

Table 9: Mobility user opinions on satisfaction

For improving the user satisfaction for the mobile application, there are two focus points for the future work. Based on the evaluation there is a need to investigate possibilities to include a feature that would enable the user to start recording the journey with less effort either by automatic starting or possibly a reminder signal. Also, another major development need for the application is to provide the user a possibility to define which messages are received automatically and which they want to be shown only when browsing through messages. Also possibility to filter messages should be considered. Even though reliability of the application is evaluated separately, it must be noted here that when aiming for good satisfaction of use, the application also needs to be more stable and reliable.

3.2.1.5 Efficiency

The efficiency of use received mixed results; some users found the application effortless to use after the learning phase whereas some experienced it to be slightly inefficient throughout the trial. Only few comments on how to improve the efficiency were received though. Some users thought it would be easier if other functionalities, such as reporting issues would be accessible straight from the recording view.

Mainly positive feedback of the application response time was received and the response time did not receive any additional comments. It must be noted that the response time is also dependent of the used device and even though improvements to response time need to

be paid attention to, it is highly possible that this part will not be possible for the efficiency to reach perfect level in all devices.

	Question	Average	0	1	2	3	4
Efficiency	Application was quickly responsive	2.8	0 %	0 %	30 %	50 %	10 %
	After I learned to use the application, using it was efficient	3	0 %	0 %	11 %	78 %	11 %

Table 10: Mobility user opinions on efficiency

From an efficiency point of view no major improvements are required based on the first trial evaluation, only some more attention to layout design in terms of positioning the navigation buttons needs to be paid.

3.2.1.6 Feedback & errors

The users stated that more feedback from the system was needed. The application did not give warnings to the users related to errors they have made. Based on users comments it would've been useful if the system had given a warning if required sensors were not enabled or there was any problems related to them. Also an acknowledgement of successful transfer of data was seen highly useful by many as no actual confirmation was received and the users had no way to know if the transfer was successful.

Regarding self-caused errors users reported having made less than 2 errors on average for every 10 times they used the application which can be stated to be rather high amount. However, when asked if they were able to avoid making the same mistakes again, the users agreed to this. No specifying comments were made regarding making same errors repeatedly, which would indicate that once the users are fully familiar with the application, errors would seldom be made.

	Question	Average	0	1	2	3	4
Feedback& Errors	Errors made by me are easy to learn from	2.2	40 %	10 %	20 %	30 %	0 %
	The application provides enough warnings if it did not function properly (e.g. due to disabled sensors)	2.7	22 %	11 %	22 %	22 %	22 %
	The application informed clearly if different actions failed	2.3	0 %	22 %	44 %	11 %	22 %
	Out of every 10 uses, how many times did you make errors?	1.5					
	I was able to avoid making the same errors again	3.4	0 %	0 %	0 %	56 %	44 %

Table 11: Mobility user opinions on feedback and errors

Based on the evaluation, the feedback-features need improvement and more communication from the application to the user needs to be developed. Special attention should be paid to providing information to the users regarding the success of different actions, such as data transfer or saving information.

3.2.1.7 Accuracy

Users have also been asked how accurate the provided information on the application is. They were asked to provide estimate of total number of times they detected the application showing a detection of activity or service line and the number of jams reported in the area

where they travelled and out of these times, how many times the detection or alert proved accurate. Many stated the activity recognition and service line detection not being reliable enough at this stage even though correct detections were also provided.

Alerts of traffic jams were received, but no experience on the accuracy was received from the users as generally they were travelling elsewhere during the alert so no reliable comments on this was received for evaluation. Therefore manual checks on data accuracy will be further evaluated within upcoming months. Even if the user experiences were not overly positive in terms of accuracy, the received feedback was expected as the modules were known to need further adjustments during the rest of the project duration and we are positive the module accuracy will reach an acceptable level for the second trial.

	Question	Average	0	1	2	3	4
Accuracy	The messages and alerts received were accurate	2.6	70 %	0 %	10 %	20 %	0%
	Accuracy of the detected activity	66%					
	Accuracy of the detected service line	65%					
	Accuracy of the jam detection	n/a					

Table 12: Mobility user opinions on accuracy

Accuracy of provided information is a critical aspect for the success of the use case. Users should be able to fully rely on the information provided by the system in order to feel at ease and be comfortable and satisfied with the use. Therefore a close attention to this part needs to be paid in the future development and the modules need to be further analyzed and improved before the second trial.

3.2.1.8 Reliability

Users were happy with the application starting reliably. However, the stability of the application was not on a level anticipated and most users experienced application crashes and unresponsiveness. In total 64 crashes were detected, meaning that of all sessions 3% resulted in a crash. Improvements and updates to the application were done during the project. Despite this crashes occurred steadily until the end. It also appears that the crashes only affected part of the user group, some users stated being very happy with the reliability whereas some reported the application crashing so often that it disturbed the use to the extent they did not want to use it any longer and felt frustrated. As the crashes occurred only with a few users, a closer evaluation of the reasons needs to be done.

Also comments on the recording problems when the application was running on the back were given, it was stated the route recorded was not precisely detected and some users stated the recording also stopping if other applications using GPS-information was used in between.

	Question	Average	0	1	2	3	4
Reliability	The application started well	3.8	0 %	0 %	0 %	30 %	70 %
	The application was stable	2.5	10 %	10 %	40 %	20 %	20 %

If you experienced crashes or unresponsiveness, how often?	9,8 times					
Application was reliable and did not crash disturbingly often	2.2	0 %	33 %	22 %	33 %	11 %

Table 13: Mobility user opinions on reliability

The reliability of the application was not optimal during the trial. Despite users mainly stating being satisfied with the reliability it appears that some devices experienced significant problems. This is a major factor for satisfaction and the success and acceptance of the service is dependent on user satisfaction. Therefore improvements for the reliability need to be made in the future to ensure user satisfaction.

3.2.1.9   Summary of usability evaluation

From the usability aspect, there are several points to focus when proceeding towards the second trial. Out of eight evaluated usability sub-categories, we found need for concrete improvements in six of them and also the remaining two still need minor modifications and adjustments for achieving even better results. The six subcategories requiring concrete actions are learnability, satisfaction, feedback, accuracy, utility and reliability.

For learnability the main point of improvement is in the user instructions and making sure that as the trial will only be running for a short period of time, the users are provided with as concrete instructions as possible so that they can fully understand the different features and utilize the whole period of the trial for the actual use of the application. Logics of the application was clear at this trial, but as more features will be included in the second, a close attention to this aspect needs also to be paid in the integration and layout planning.

The improvement needs for user satisfaction are threefold. The possibility to need less effort to start recording the journey needs consideration. Also, the users should be provided a possibility to define if they want to receive messages automatically or not and possibly even have the chance to determine which messages are received automatically and which they want to be shown only when browsing through messages. Also possibility to filter messages should be considered.

Major improvements for the feedback from the system are also needed. This was a category left with minor attention in the setup but it is clearly noted that more information regarding the application functions need to be given to the user. Special attention should be paid to providing information to the users regarding the success of different actions, such as data transfer or saving information and also more warnings of problems with the application, such as disabled sensors, should be provided.

The accuracy of the provided information is a category where improvements need to be accomplished. This is a critical factor for the users in both acceptance and satisfaction and therefore one of the most crucial factors for the success of the whole service and needs to be paid close attention.

In terms of utility, we are satisfied with the results, major points of improvements were detected based on the user questionnaires, but all of them were those already included in the primary plans for the second field trial and were mainly about provisioning of more information to the users. However, the user also needs to be able to better have a possibility to filter unwanted information, so that the feeling of usefulness is greater when unnecessary frustration with useless information flood can be avoided.

Also the reliability of the application needs more focus, even though improvements in the application were made during the trial some users had critical difficulties with using the application due to crashes. The problems were focused only to some users whereas with some devices the application functioned without problems and therefore the reasons behind problems with certain devices need to be investigated and solved.

3.2.2   Government web services

As part of the trial, the HSL contact group for Live+Gov were asked to give their impressions on the issue reporting portal and the web interface for recorded routes both at the current state and based on the plans for future development. An evaluation meeting was organized after the trial with four members of HSL personnel. As the focus of the first trial was more on the citizen side services, the evaluation was done with the emphasis on how to proceed with the development tasks in order to be able to provide the authorities useful tools. The provided tools were developed in order to be able to provide basic visualization of the events during the trial and to further define the desired visualization techniques for the second trial.

Due to the reasons mentioned above, the focus on the evaluation has been on defining if the development of the project is focusing on providing useful tools and if the requirements and plans are seen useful. Also the usability of the already provided tools and the impact of the trial are covered.

The general impression of the developed tools was positive. It was agreed on that the system provides much valuable information and improves the possibilities to increase interaction and dialogue with the citizens. The issue reporting feature was seen as a useful tool and to bring additional value to existing feedback-channels especially through the attached location information and the possibility to attach images. However, it was seen that since a wider feedback-channel already exists, at least for HSL it would be inefficient and uncomfortable to use two similar channels simultaneously. To avoid the situation where multiple feedback channels are used simultaneously and to make the use of provided information more efficient, a separate interface for connecting the provided information channel to the already existing systems would be needed. With the possibility to integrate the system would ensure the full benefits to be gained as well as to maximize the satisfaction level of using the system. Obviously, if no similar system already existed, the separate web service would be of much higher value, therefore this feedback is strongly HSL oriented and does not comply with all public transport authorities in Finland or within Europe.

Also the possibility to send messages and alerts directly to the application through a web interface was seen profitable and to increase interaction. The direct distribution of exception info acts as a good example of how to integrate existing services to the system. The impression was that this functionality could be taken further by enabling the possibility to filter to which users the messages would be sent and also by enabling the users to define which messages and alerts they want to receive.

The reporting view of travelled journeys is seen highly important for supporting public transport planning processes as well as the general transport infrastructure planning. The current state was already seen useful and with the planned development in the presentation and filtering of shown data was seen highly positive. It was stated that this type of

information that will be provided is the type of information planners need the most, but until now there has not been adequate and effortless use to acquire this specific travel data in the planned extent. Therefore, the travel data and possibility to visualize it as planned would increase the planning efficiency greatly. At this point, however, it was emphasized that the data quality and accuracy of detections need to be of reliable for this type of information to be used and that enough users need to be involved before the data has a meaning. Therefore it was also questioned how to acquire adequate amount of users who would use this kind of application and it was addressed that the benefits for the user need to be clearly presented.

Wider impact than only for public transport is also seen possible with the planned service. The only point of concern is how to ensure the even quality and good cross-section of citizens to use the system. A good cross section of citizens as users would be needed for enabling the acquisition of realistic data that would represent all types of citizen's preferences instead of only information of a specific type of users. It was seen that if only a narrow take of citizens was using the service, this would easily distort the interpretations made based on the received information.

When discussing the usability and utility of the service for planning and interaction purposes, there was one special focus point that was addressed as the most important factor defining the success of the service. The connectivity to the possibly existing systems needs to be taken into notice. Many municipalities already have different information systems where these types of services bring significant additional value, but if the information cannot be accessed from the existing systems, the use might become difficult to the extent that the satisfaction and interest to use it suffers. However, it was agreed that with the project is proceeding to the right direction and that during the trial it is acceptable to have separate services but if the service was to be taken to production; the integration possibility would be anticipated.

Another point that was addressed was the question of what, from the citizen perspective, makes the service so much better than all other existing services that enough citizens would use the system. Currently there are several open source application that provide the users information of the traffic situation and conditions that unless clearly marketed and emphasized there is a risk the benefits might not be easily understood and the system would be easily left unused. It was not questioned if there are enough benefits to the user but it was questioned if the users would understand those easily. Therefore when defining the exploitation plans the marketing aspect needs to be addressed also.

3.2.3   Evaluation of the requirements

3.2.3.1   Set requirements

Based on the trial we can state that the customization process has resulted with positive outcome. As stated before, from all requirements set for the use case, the most important and the most core requirements were selected for the first trial to create a realistic and important basis for the next steps of the project.

The first requirement was that *the system is able to provide communicative tools for government and citizens including reactions, alerts and messages* (L+G.FR.05). This requirement was met with the following features:

- issue reporting

- jam detection alerts
- exception information messages
- authority messaging

Multiple communication channels were implemented in the prototype. These features include communication between citizens and authorities in both ways. Therefore we can state this requirement to have been met successfully.

The second requirement was *to be able to capture images* (L+G.FR.07). This requirement was met with the issue reporting feature, where users could attach images to their issue report. AR has not been included in this trial and therefore the content recognition has not been implemented in this feature, but the requirement in terms of the use case has been successfully met.

The third requirement was *to be able to utilise 'citizens as sensors' for their environment by mining relevant data (e.g. location, context, pictures, and reports) from their reality* (L+G.FR.10). This requirement was targeted by collecting journey data and issue reports from citizens. From both, more advanced information is extracted with data analysis and visualization methods and therefore this requirement has been met successfully.

The fourth requirement stated the system having *to be able to connect to different municipality administration systems* (L+G.FR.13). Targeted for this requirement, the prototype included connection to the real-time information system from where information was collected and utilized in service line detection, and traffic jam detection. Also, the disruption info messages were obtained by communicating with an external system maintained by HSL and shown via the application to the users. With two different systems successfully connected, we can state also this requirement to be met in the trial.

The fifth requirement stated that the system *shall receive data from the application containing travel information and means of travel* (L+G.FR.14). The prototype collects movement and journey data recorded by the users based on location and time. Also specific components for detecting activity (HAR) and used service line (SLD) were implemented. With these features, the prototype contains several points that together meet the given requirement. Thus, this requirement is met in the first trial.

The last requirement was the system *to be able to detect traffic jams* (L+G.FR.17). Based on this requirement, a specific traffic jam detection module was developed and implemented. During the trial, warnings of jams have been generated by the module and the system has processed warnings to the users. With this, also the final requirement has been met.

All requirements for the first trial were successfully met. However, it must be noted that many mentioned features still require improvements and modifications for the second trial for better quality and user satisfaction but as for the current phase of the project the requirements are met in the anticipated level.

3.2.3.2 Completeness of the requirements

Even though all given requirements were met in the customization process we also need to ask if the given requirements were realistic for real life scenarios and are there more requirements that should have been set. As stated, we were able to meet all requirements on a satisfying level and the requirements given were of a general nature, formed around the aim of providing tools that would ease users with events occurring in everyday life. These events would be something that can happen anytime, anywhere in urban

surroundings. We see that any tool that can ease a citizen's life or improve development actions on authority side are useful and if tools can be developed the requirement has been proved realistic. The general opinion was that these requirements are realistic and therefore we see that no major modifications need to be made for the final trial.

From authority point of view the set requirements stand reasonable and they have been met in an acceptable level. However, it was clear that most features still require more development in order to be of true value for the authority side in general and also in order to form a level of usability so that the service is comfortable to use. The most important aspect to pay further focus on from the authority point of view is the integration possibilities to existing systems in greater extent as well as to deepen the reporting side.

As all requirements were met on a satisfying level we are satisfied with the customization process and can state the prototype delivering all aspects expected to be involved in the first trial. The feedback from the trial users has been valuable during the trial and several points of improvement have been taken into account when starting the process towards the following trial.

Another interesting point of view for evaluating the requirements is to ask if all necessary requirements were defined or were there any missing requirements that should have been set for the trial. As we will later see, the results of the trial have been good and the trial successful and no severe insufficiencies can be stated to have been resulted in terms of requirements. However, there are some factors that would have led to even better results only if we had known to include these parts to requirements. In this trial we have mainly focused on the technical parts, aiming to create a combination of different modules that work together as aimed, but the softer user experience point of view was not taken into account in a detailed way in requirements. Even though the user experience was paid some attention to during customisation phase, more emphasis on this aspect should have been included in the very first steps of the trial setup.

3.2.4   Evaluation results for specific WPs

In this section, we present specific input for WP1-WP4, based on our evaluations.

3.2.4.1   WP1 – Reality Sensing and Mining

Data quality

In total in the trial we gathered more than 85 hours of travel data from the users. Based on the HAR and SLD data the first impression of the data quality seems plausible for the trial success. However, more detailed analysis of the data and the content needs to be done. The number of sessions in total during the trial was 1,784, out of these 96% of the use was done in the defined trial area which covered all of HSL service area with the following distribution:
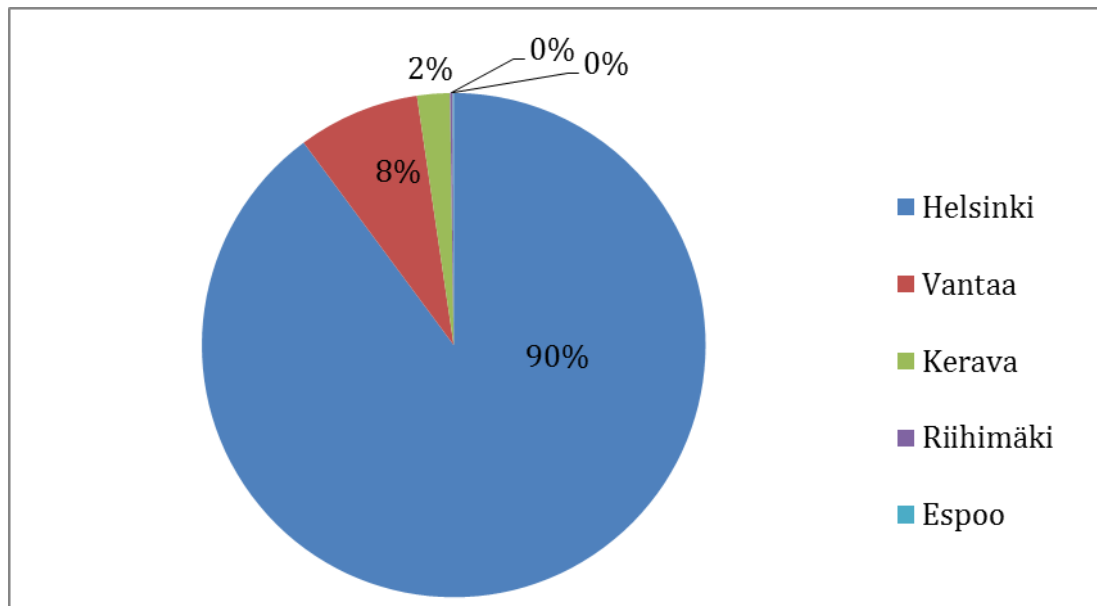
Figure 3: Distribution of sessions in the trial area

The sessions outside of HSL area mainly consisted of those done in Thessaloniki or Tampere by the Live+Gov project members when checking on the application. Also a few additional sessions were detected outside the trial area, which can mainly be explained by trial users travelling around Finland. These external sessions not done by project members included only 1.2% of all sessions, therefore we are satisfied with the geographical coverage of the trial use and are happy to confirm no distorting use was detected.

The human activity recognition (HAR) accuracy on the trial evaluation was also done based on comparing the manual annotations by the users with the automatically detected activities. The results show that during the trial 222,278 samples for HAR were collected in total. The most popular activity would have been standing, which included 75% of all gathered results, which, at this stage, seems plausible but the analysis needs to be made in more detail to ensure the quality. In Figure 3 we have presented the distribution of detected activities.
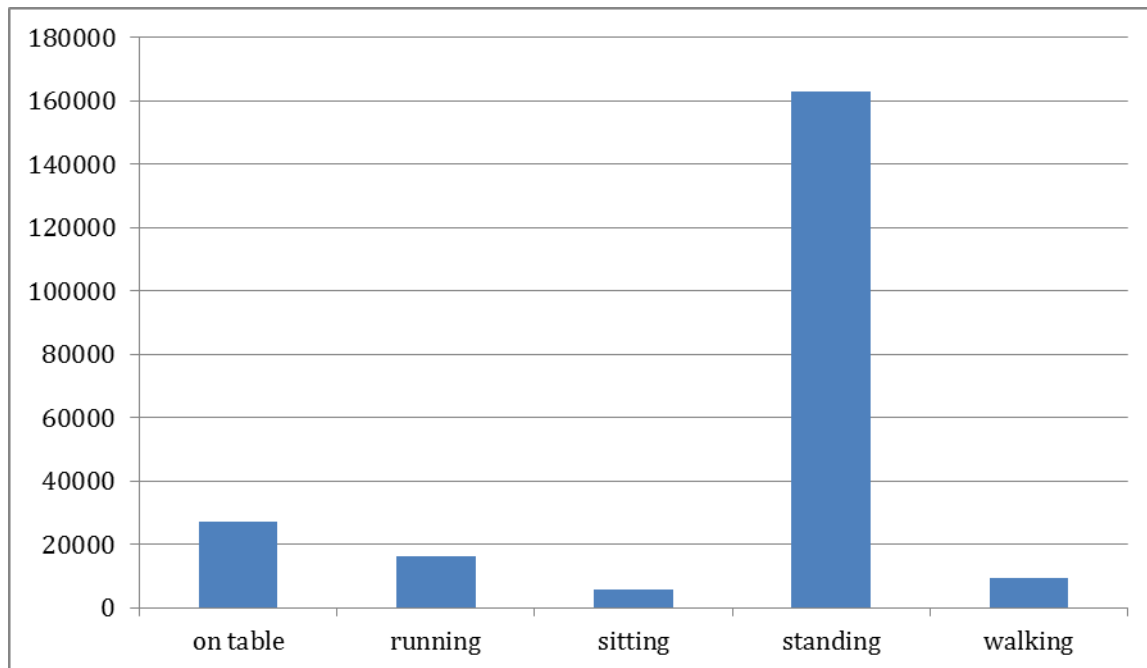
Figure 4: HAR activities detected during the Mobility trial

When asked from the trial users the number of times they noticed the application detecting the activity, how many times the detection was accurate and how many times they corrected the false detection the following results were received: Out of all detections 66% were estimated to be correct and out of all false detections 32% were corrected. This estimate gives us a good impression of the accuracy of HAR, but more detailed evaluation of the accuracy needs to be conducted separately.

Service line detection accuracy on the trial evaluation is mainly done based on comparing the manual annotations by the users with the automatically detected lines. The results show that during the trial all together 2,327 times the API was called in Helsinki region. From these there were 133 routes that were manually tagged. Out of all the manually tagged routes, 22 were detected correctly. This would mean the accuracy of the detection being 16.5%. However at this point we must emphasise that the accuracy of the detection is most likely much higher as the users generally do not re-tag a correctly detected route.

This primary impression of accuracy is supported by the user experiences. When asked from the trial users the number of times they noticed the application detecting the service line, how many times the detection was accurate and how many times they corrected the false detection the following results were received: Out of all detections 65% were estimated to be correct and out of all false detections 30% were corrected. The data gathered and detections done need to be analysed in more detail as based on the pure statistics no realistic impression can be received.

The traffic jam detection worked as planned during the trial. Logs of the detected jams were implemented during the trial for evaluation purposes for both accuracy and stability. Based on the logs, the detection rate has been stable most of the time, every week during the trial jams have been detected and only few times the API has been noticed to be inactive due to server-side problems either on the module side or the public transport system side. In the trial parameters were set moderately tight so that as little false alerts would be generated as possible but still jams would be detected regularly.

In total alerts regarding traffic jams were given on average 3 times for each session during the trial, resulting in slightly greater number of detections than anticipated, which would indicate the parameters needing to be set tighter in the future if only greater jams are wanted to be detected. However, despite the number of jams detected users reported never actually travelling during the alert on the given area. Also, during the trial no severe jams are known to have happened in the covered area so manual checks could not have been made and the accuracy cannot be verified at this stage. Due to these reasons, manual comparison to other known services providing jam alerts was done to provide an understanding of the accuracy of the detection. The comparison showed similar detections, but in all cases the detected jams were of a short duration, mainly less than a minute and therefore no reliable statement about accuracy could be done based on this information at this stage. However, the module will be further tested after the trial to ensure accuracy for the final trial and the parameters will be re-defined after the tests. Also the possibility to combine information from other existing services needs to be closer examined together with an evaluation of usefulness of including the users a possibility to report jams when they travel.

3.2.4.2   WP2 – Policy Modelling and Visualization for Mobile-aware Dialogue

The most interesting results for WP2 were gathered from the evaluation meeting in terms of future development of the service. It was noted that as such, the application did not create much impact in the authority processes but there was some interaction already created towards citizens in terms of presenting information coming from SLD, TJD and HAR, as well as allowing the user to make corrections to these. Also, by providing users with alerts and messages and allowing them to report issues, we were able to generate a basic form of citizen-government dialogue.

For the visualization it is clear the outcome needs to be clearly visible and information acquisition needs to be effortless. This would mean all information sources needing to be accessible from one source and in understandable format.

Also, for the second trial we need to focus more on the trial setting also in the authority side. More preparations before the trial need to be made in terms of engaging authorities, providing instructions and guidance for using the tools and ensuring the participation level remains on anticipated level throughout the trial in order to create real impact in the authority side. The aspects, which will be covered in the evaluation of the second field trial, will be extended in reference to the planning process as described D2.3 [12]. There, a clear advantage of the Live+Gov system has been identified as compared to the current system of planning the traffic infrastructure. This advantage, which will be tested and evaluated, relates to the accuracy and the cost efficiency of the planning process.

3.2.4.3   WP3 – Mobile Augmented Reality

The most significant results for WP3 on this trial evaluation were received regarding the user experience and acceptance on personalized information. When asking the user opinion on the usefulness regarding the personalized information during travel, it was stated having a positive impact. Users agreed on finding the personalized content useful and pleasing to have. It was stated that the application provided a useful tool for personal travel and that the personalized aspect made it both more interesting and more useful.

What is needed to develop further in the future work is to deepen the level of personalization in the information. The trial users commented on finding it useful to include prediction of user movements to further filter the information as some users strongly stated

that there was too much general information that was not relevant to them and that the big amount of information received resulted in the situation where the relevant information was easily left unnoticed. An important factor for the personalized information acceptance is also the accuracy of the information provided by components in WP1 and the users need to be able to rely on the information before accepting it.

3.2.4.4 WP4 – System Integration

During the trial all together 15 users tested the application with their smartphones and four devices were also used by the project members to verify the application functionality. Based on the comments and observations the overall impression is that the different components are working together well. No major issues occurred during the trial. Minor issues remained during the trial, e.g. problems with missing photographs and alert distribution occurred. However, these issues were not obstructing trial execution and most of them were successfully solved in intermediate updates of the trial components.

During the trials mixed signals were received regarding the stability of the mobile application. Some users reported no problems at all, while other users suffered several crashes, which not could be solved during the trial.

Stability needs more focus on the second trial in terms of component and system tests. Even though the application was tested in detail by project members before the launch with positive results, it appears that tests done with more variety of devices and android versions could have helped to better detect problems before or during the trial. It is also clear that due to short development cycles, the full needed attention was not possible to give in the integration and testing thoroughly and with the variety that would have been required to provide fully reliable application.

3.2.5 Impact and success of the Use Case Trial

What was learned in the setup of the first trial was that the unexpected challenges in the user recruitment need to be strongly addressed in the second trial setup. The second trial will be organized for a significantly larger group of users. This means that different recruitment channels need to be closely evaluated when planning the recruitment process. Also, based on the first trial experiences, the recruitment needs to be started significantly earlier than in the first trial in order to have better reaction time for unexpected development in the setup process. For the second trial more interaction with the authorities is also needed throughout the trial and therefore engagement of the authorities needs significantly more effort and attention to be paid than in the first trial.

With the wider coverage, bigger user groups and the planned development the impact of the application also increases significantly. At this point, the citizen impact was remained rather small in many parts but we are confident that with the planned actions for the second trial, we will be able to create a service that will have a much stronger impact for the citizens.

For the final trial in the Mobility use case, more media coverage is also anticipated. As the first trial was intended to be done within a closed user group, it was also intended not to aim for media coverage. However, as the final trial is intended to include a wider user group and a more comprehensive service, it is highly intended to aim for coverage in media. For this, a detailed plan should be included when setting up the next trial.

Positive feedback of the application and the planned development activities was received by the users and user activity rates were acceptable during the trial. However, as expected, it was clear that the users did not use the application constantly and that not all features were fully utilized in the use. We see this as a result from the fact that the service is not yet as comprehensive as it will be, many features still lacking from the prototype, which leads to the situation where only minor benefits are yet offered. Therefore the future work is well needed for gaining more interest to use the service. Creating more interest for the use amongst citizens is a critical factor for the success of the developed services and therefore we need to focus even more on developing ways to engage and activate users even more.

By engaging the participants also better usage numbers will increase and more data will be gathered. At this point the limited number of users provided enough data for the trial purposes, but when aiming at a full scale service where authorities are meant to be provided a comprehensive understanding of citizen movement and improvement needs of the services, no decisions can be made with such a small number of users or data. Therefore a wider coverage of the application will be needed and in the second trial we need to aim at a wider coverage user number wise.

Based on user activity, data quality and the customization process the trial has provided satisfying results. Many aspects of the application need improvements for the final trial but the work is progressing steadily and to the right direction. General comments from users were various and the work is seen interesting and useful by many users.

3.3   Summary of the evaluation of Mobility

General opinions of the usefulness and impact of the trial for citizen life were relatively wide spread among users. Based on the comments received, those of the users who took into notice the planned development activities of the service thought that the impact will be significant. Some users lacked the vision of the development plans and stated that the prototype at its current stage would not have enough impact for them to use the service in the future. We see this as a positive result. The development stage of the prototype is not yet final and many features the users stated being useful for improving impact and the satisfaction of use are planned to be implemented.

When asking the users if they regarded the application as a useful tool and as a useful information provider the general feedback was positive. The application was generally seen useful for the public transport user and the interaction possibilities were regarded both useful and interesting. However, during the trial interaction was still limited. Even though the application provided a tool for interaction only few users reported any issues during the trial and no specific information was yet provided to the users by the authorities so the impact remained small in the trial. This is an issue that needs to be addressed and more ways to engage both citizens and authorities in the next trial need to be considered.

After the trial the users were also asked how they see the future plans of the use case and whether they see these plans resulting in a service they would be pleased to use. A description of the plans was given and the users were asked specific questions regarding their thoughts of the usefulness of different features.

The questions presented to the users were about if they would find it interesting to be able to filter alerts and messages and to receive alerts on the lines they are currently travelling on, if they would find it useful to receive alerts on the lines they travel often with before journey even started, or to receive feedback on the reported issues they have provided. Finally they were also asked if they were interested in receiving information regarding planned changes in public transport services and commenting them. All of these mentioned features were found useful and interesting, only the possibility to comment planned changes in public transport received mixed feedback. Those who thought it would not be interesting feature in the application stated that it would be enough if the application only informed about new plans and provided a link to external page where also comments could be given. Also, what was pleasing, highly positive comments about the future plans and development so far were received by the users despite the detected problems.

At the end of the evaluation of future work, users were also asked if they would like to participate also in the next trial and general opinion was positive. Therefore we are assured the future plans for this use case are regarded as useful from the citizen perspective and in general the citizens' opinion was that the development is heading to the right direction.

The HSL representative's opinion on the developed tools and the progress in general has been positive. It is clear that more development is expected before we reach our goals and there are still issues that need to be paid close attention to, but the progress is satisfying and plans are seen both realistic and useful. With the first trial evaluation more valuable information and insight on the authority side requirements have been received, making future development more clear and detailed.

As a result we can say that as expected the impact of the trial has been rather low for public, whereas for the project significantly high. As for the following trial major points of improvement for the external impact include actions for getting more media coverage and improving the communication between citizens and authorities and between citizens and the application. At this stage the required tools exist but most features require improvements to various parts. Also, for a stronger impact in the second trial we need to focus on finding efficient ways to engage the users.

4   Urban Maintenance Use Case: *Jij Maakt Utrecht*

In this section we present the evaluation of the first field trial of *Jij Maakt Utrecht* for Urban Maintenance. Following the general evaluation strategy of WP5, we present results on the usability of our customised prototype, evaluate the requirements, give specific input for WP1-WP4 and describe the Use Case impact. Before we present our results (in section 4.2), we will describe the field-trial set-up, give a brief overview of the used prototype, and the tailored evaluation process in section 4.1). In Section 4.3 we will present a global overview of the *Jij Maakt Utrecht* field trial and its results.

4.1   *Jij Maakt Utrecht* Field Trial

4.1.1   Field trial set-up

Jij Maakt Utrecht is the customised version of the webapplication for eGovernment Dialogue and Visualisation and is described in great detail in *D5.2 – Prototypes-demonstrators for first trials*. The webapplication *Jij Maakt Utrecht* has several goals that have been described in D5.1 for its initial (functional) design. Overall, it provides digital support for citizens' participation and collaboration initiatives within the municipality of Utrecht. In other words, it supports all three pillars of citizen participation: *Transparency* by showing relevant information, *Public participation* by supporting governments in actively asking input from citizens on certain topics, and *Collaboration* by supporting co-maintenance of a certain object in public space [5]. To achieve this, there are four sub-goals, defined from a citizen perspective:

1. *What is going on in my neighbourhood?* The webapplication addresses the information needs of citizens.
2. *Where can I add influence?* The webapplication supports policy- and decision-making: participation.
3. *What can I undertake myself and/or what can I organise with others?* The webapplication responds to the self-organising ability of neighbourhoods and areas: collaboration.
4. *Which data/tools/applications are available, can I use for other applications, what can I share myself and/or with others, what can I add?* The Webapplication prompts local economy and engagement by making open data sets available to citizens.[1]

The webapplication *Jij Maakt Utrecht* is developed in co-creation with the municipality of Utrecht and other stakeholders. Characteristic for a co-creation process is that it is an undefined process, dependent of the participants' [16]. In organisational management theories, co-creation is related to the creation of economic value by working together with consumers [17]. It is an innovative strategy where designers and product developers involve consumers in the design process. Governments in the Netherlands are looking for ways to use co-creation to involve citizens in the policy process. The process of co-creation covers three ideas [18]:

1. Creating something new
2. Working together on a basis of equality

And the third idea that follows from 1 and 2:

3. 'change to work on change': while working on transformation, the organisation itself simultaneously transforms as well.

Co-creation is not a predetermined process and the end result is not predefined upfront. Moreover, when increasing the level of participation in the process, the level of structure will decrease.[18] Therefore, one cannot say that there were clearly defined development stages in this development of *Jij Maakt Utrecht*. Design, collecting requirements, customisation and evaluation are done simultaneously during the entire process. At the start of the co-creation process of *Jij Maakt Utrecht*, the focus is on collecting requirements, on the design and on the customisation of the webapplication. The goal was to release a working prototype that could be trialled in a public setting on May 30, 2013. After the start of the public part of the trial, the focus shifts more to evaluation, but design and collecting requirements still remain important. This process is visualised in Figure 5.
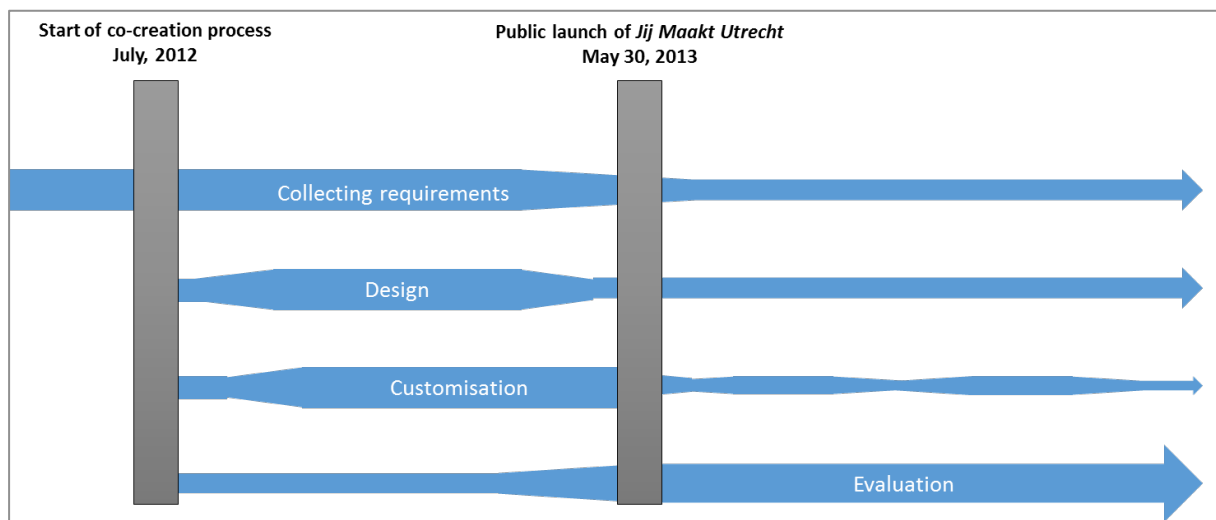
Figure 5: Stages of Collecting requirements, Design, Customisation and Evaluation in the Co-creation process of *Jij Maakt Utrecht*

When the evaluation results for the first trial cycle are reported, evaluation for *Jij Maakt Utrecht* does not end. In the next steps, focus will shift again to refining the requirements and design, while customisation continues to take place.

4.1.2   *Jij Maakt Utrecht* prototype

The first *Jij Maakt Utrecht* trial ought to fulfil the Live+Gov functional requirements as shown in Table 14 that were assessed in D5.1. [2]

#	Live+Gov functional requirement
L+G-FR.03	Be able to show information in a context aware manner (location, person, time-specific, situational) to create greater awareness and create mutual citizen-government understanding
L+G-FR.04	Provide possibilities to group, aggregate, filter large amounts of data to discover knowledge
L+G-FR.05	Provide communicative tools for government and citizens (including reactions, alerts, messages)
L+G-FR.06	Provide possibilities to visualize knowledge and patterns from data by data-mining techniques

L+G-FR.13	Be able to connect to different municipality administration systems
L+G-FR.20	Provide possibilities to present co-maintenance initiatives in an area
L+G-FR.25	Improve the quality of existing (open) data sets related to urban space, by leverage of citizens functioning as sensors and verify with existing registrations
L+G-FR.26	Introduce elements of serious gaming / gamification

Table 14: Requirements for *Jij Maakt Utrecht*

In Table 15 the involved Live+Gov components for this trial are summarised from D5.2 [1]:

#	Live+Gov component	Type	Description
C1	Issue Reporting Service	Back-end web services	Issue Reporting (BuitenBeter) Service. Back-end Web Service for the mobile data-collection prototype, which allows filing a report.
C2	Issue Distribution Service	Back-end web services	The server side service for distributing the specific data collected with the mobile data-collection prototype. In this case, the reports are filed to a separate server.
C4	eGovernment Dialogue and Visualisation Service	Back-end web services	Service responsible for storage of initiatives, feedback, and photo's for the Urban Maintenance specific prototype *Jij Maakt Utrecht*.
C11	Mobile Issue Reporting Client	Mobile Application	Customised version of BuitenBeter. Specific purpose of gathering data as input for *Jij Maakt Utrecht*.
C18	Webapplication for eGovernment Dialogue and Visualisation	Webapplication	Webapplication responsible for visualisations and interaction (participation, collaboration and initiatives) between government officials and citizens via the browser. The customised version is *Jij Maakt Utrecht*: the specific webapplication for the Urban Maintenance Use Case.

Table 15: *Jij Maakt Utrecht* including mobile data-collection prototype: Live+Gov components

4.1.3   Evaluation process

As described in 4.1.1 and visualised in Figure 5, for *Jij Maakt Utrecht*, the process of developing, trialling and evaluation is not completely linear. The philosophy for the Use Case is coherent with this: public beta release of the application is not aimed to be the end-result; it is aimed to make co-creation with users (citizens and officials) possible. This, by nature, means that there are multiple evaluation moments during the trial and that there is also

scope shifting during the trial. The prototype is a process catalyser to form ideas and strategies on participation and urban space.

For the evaluation, the results are gathered using quantitative and qualitative research methods. Collecting qualitative data for evaluation purposes has been done during the entire co-creation process. Some important evaluation results are directly implemented, because of the need to release a working prototype for the beta version on May 30, 2013. The collection of quantitative data is mainly done during the public part of the trial.

4.1.3.1 Information sources for evaluation

Qualitative research methods are used during the entire co-creation process to collect information for evaluation purposes. Linked to the unstructured nature of the co-creation process, most of the data is collected in a rather unstructured way. This is done in many interaction moments during the trial period: several LEAN-sessions, co-creation sessions, meeting with public officials, stakeholders and users as well as a meeting with the mayor of Mundaka (Spain), a Council information evening, events where *Jij Maakt Utrecht* was shown, etc. The evaluation information sources for *Jij Maakt Utrecht* include the list as presented in Table 16.

Date	Methodology
13 October 2012	Prototype evaluation meeting with stakeholders municipality Utrecht
29 November 2012	Observational collection of initial user impressions
20 March 2013	Interview with leading stakeholder (manager participation) municipality Utrecht
22 March 2013	Status evaluation meeting including information manager municipality Utrecht
27 March 2013	Status evaluation meeting, including urban designer and participation manager municipality Utrecht
11 April 2013	Evaluation meeting for integration of specific serious gaming elements
6 May 2013	Evaluation acceptance meeting for public launch
6-17 May 2013	Direct observation within neighbourhood departments of the municipality 5 unstructured interviews
7 May 2013	Interviews with two public officials on tactical level: a neighbourhood manager and public official with the City Maintenance department
30 May 2013	Observational collection of initial user impressions
1 June 2013	Behavioural observation of 25 citizens interacting with the system
1 July 2013	LEAN session 1 municipality Utrecht - working processes
9 July 2013	LEAN session 2 municipality Utrecht - ICT & Infrastructure
17 October 2013	Evaluation meeting municipality Utrecht including management module
22 October 2013	Interview with content manager municipality Utrecht
11 November 2013	Short interview with program manager implementation open data in Utrecht
22 November 2013	Meeting with Aitor Egurrola, mayor of Mundaka (Spain) about *Jij Maakt Utrecht* possibilities
13 December 2013	Evaluation meeting municipality Utrecht
10 January 2014	Targeted evaluation meeting with neighbourhood councils in Utrecht
16 January 2014	End survey questionnaire for *Jij Maakt Utrecht* digitally published

17 January 2014	Initiative Market: Behavioural observation of about 15 citizens interacting with the system Distribution of questionnaires to the event participants
20 January 2014	Co-creation evaluation meeting and survey questionnaires for key-group *Jij Maakt Utrecht*
23 January 2014	LEAN session 3 municipality Utrecht - Management & Control
27 January 2014	LEAN session 4 municipality Utrecht - People & Culture
28 January 2014	Evaluation session for public officials about the 'Initiative Compass'
28 January 2014	Council Information evening session on initiatives in *Jij Maakt Utrecht*

Table 16: Evaluation information sources for *Jij Maakt Utrecht*

In general, these evaluation information sources cover the following levels:

- Citizens   (impression in Figure 6)
- Stakeholders   (impression in Figure 7)
- Public administration organisational   (impression in Figure 8)
- Political   (impression in Figure 9)

Figure 6: Initiative market impression *Jij Maakt Utrecht:* citizen level

Figure 7: Co-creation meeting impression *Jij Maakt Utrecht:* stakeholder level

Figure 8: Learn session impression *Jij Maakt Utrecht*: organisational level

Figure 9: Council meeting impression *Jij Maakt Utrecht*: political level

Apart of these more qualitative sources, we also exploited the more quantitative methods;

- Distributed surveys to registered users, a general survey published on the webapplication and paper surveys brought to meetings.
    - Public officials
    - Citizens - initiators
- Data about usage of *Jij Maakt Utrecht* from the database and web analytics

The following Table 17 summarises the evaluation aspects for this use case on a high-level. For each aspect the used evaluation methods are given as well as the the WP the evaluation of the specific aspect is delivering results to.

Prototype evaluation				
Aspect	**Criteria**	**Specifications**	**Targeted WP**	**Methods used**
Customization process success	Integration	How well were included components integrated?	WP4 & WP5	Internal review
	Compatibility	How well does the application	WP4	Internal review

		perform on different platforms		
	Stability	How well does the application performs with big-data	WP4 & WP5	Internal review, data analysis, log files
Data mining & knowledge discovery	Pattern recognition	Does the application discover knowledge patterns	WP1 & WP2	Interview sessions, discussions, internal review
Usability	Learnability	How easy is it for users to accomplish basic tasks the first time they encounter the design?	WP5 & WP2	questionnaires, personal contact
	Utility	Does the prototype meet the given requirements	WP5 & WP2	Evaluation sessions, questionnaires, internal review
	Memorability	When users return to the design after a period of not using it, how easily can they re-establish proficiency?	WP5 & WP2	questionnaires, personal contact
	Satisfaction	How pleasant is it to use the system in general	WP5 & WP2	questionnaires, personal contact
	Efficiency	Once users have learned the design, how quickly can they perform tasks?	WP5 & WP2	questionnaires, personal contact
	Feedback & Errors	How many errors do users make, how severe are these errors, and how easily	WP5 & WP4	questionnaires, personal contact, data analysis

		can they recover from the errors.		
	Feedback & Errors	Feedback from the system to user	WP5 & WP2	questionnaires, personal contact, data analysis
	Reliability	Does the software runs consistently without crashing	WP5 & WP4	questionnaires, personal contact, data analysis
Visualisation tool	Collaboration	Does the webapplication succeed in creating collaborations?	WP2	Interview sessions, discussions, internal review
	Participation	Does the webapplication succeed in creating participation?	WP2	Interview sessions, discussions, internal review
	Transparency	Does the webapplication succeed in creating transparency?	WP2	Interview sessions, discussions, internal review
Impact	General	User numbers	WP5 & WP4	Data analysis
		Usage numbers	WP5 & WP4	Data analysis
		Media coverage	WP5 & WP6	Internal review
	Embedding in organisation	Is the system embedded in the organizations and adapted in processes, and policies	WP2 & WP5	Interview sessions, discussions, internal review

Table 17: Evaluation aspects and methods in the first *Jij Maakt Utrecht* trial

4.2   User-experiences and customised prototype evaluation

In this section we present the results of the first field trial of *Jij Maakt Utrecht*. First in 4.2.1 the evaluation of the usability of the webapplication is given. In 4.2.2 we will focus on the requirements as set out in D5.1, and check how the requirements for the first trial are experienced. Finally, in 4.2.3 we give the specific input for WP1-WP4 for the next phase of Live+Gov. All the evaluations are done from a functional user perspective.

4.2.1   Evaluation of Usability

Usability is measured by a total of 18 questions, divided over seven topics. The 43 respondents rated to what extent they agree with a statement, on a Likert-scale of five levels:

1. strongly disagree
2. disagree
3. neutral
4. agree
5. strongly agree

The average score on usability is 3.33, which is between the levels 'neutral' and 'agree' (see Figure 10).
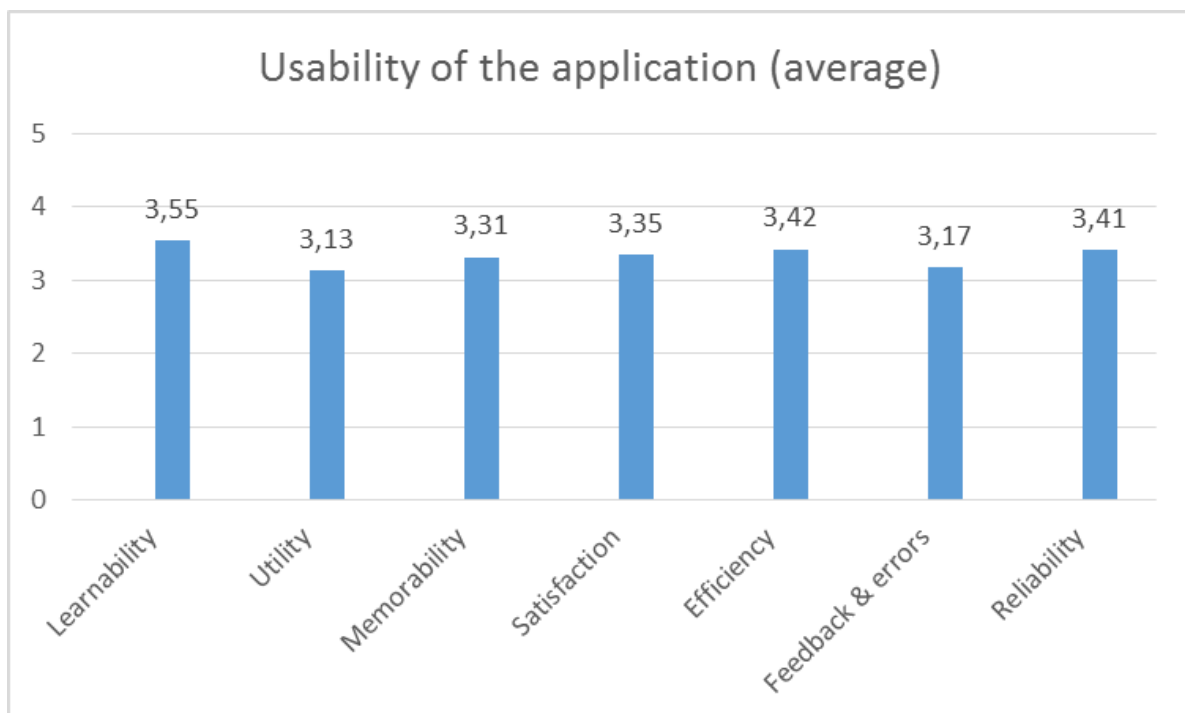
Figure 10: Average score on usability topics for *Jij Maakt Utrecht*

In the following sub-sections, each topic will be described separately. In these sections the following terms are used:

- Mean (μ): average score
- Median: the median is the 'middle value': 50 % of the scores is less, 50% of the scores is more than the median

- Standard Deviation (σ): by normal distribution: two-third of the scores is between μ-σ and μ+σ, and 95% between μ-2σ and μ+2σ – says something about distribution of scores, the higher the more distributed.

4.2.1.1 Learnability

Learnability is measured along different aspects of Learnability. In following Table 18, the results of Learnability and its sub-aspects is given.

Question topic	Mean	Median	Standard Deviation
Learnability overall	3.55	4	0.93
Easy to learn	3.72	4	0.89
Easy to find information	3.41	4	0.91
Easy to navigate	3.38	3	1.04
Use without written instructions	3.78	4	1.01
Clear language	3.53	3.5	0.84
Intuitive	3.47	4	0.88

Table 18: Descriptive statistics for Learnability for *Jij Maakt Utrecht* webapplication

Per aspect, we analysed the information to see what specific lessons could be learned for the learnability of *Jij Maakt Utrecht*.

These scores show clearly that on the sub-aspects of the Clearness of Language and the Ease to Navigate, the scores leave room for improvement. The other topics show positive scores.

The results show quite clearly that the webapplication itself is easy to master without the use of any manual or instruction and that it doesn't rely on a high learning curve. However, the content of the application and the different purposes and goals of *Jij Maakt Utrecht* seem to make the results lean towards a more neutral user experience.

This user-experience impression is also strengthened by the information of our evaluation meetings and interviews. The purpose and goals for *Jij Maakt Utrecht* have shifted along the process for the municipality. Also, the first sprint focussed on publishing material to see what is possible. Along the process, the concepts and goals for *Jij Maakt Utrecht* have been subject of discussion, mostly internally for the stakeholders and on the organisational level for the municipality. This is in accordance with the chosen approach: the public beta is the meant to gain insights and start the process. Overall, it is not totally surprising that the scores there are very positive results on the easiness to learn to use the application and a bit more neutral opinions regarding the language and accompanied navigation.

4.2.1.2 Utility

Important aspect for the usability is that the system actually follows the expectations of the users regarding what can be done with the application. This aspect is described in greater detail in 4.2.1, where we evaluate the functional requirements. In our questionnaire surveys, we found the following user-experience on "I can do what I expect with *Jij Maakt Utrecht*".

Question topic	Mean	Median	Standard Deviation
Utility overall	3.13	3	1.04

Table 19: Descriptive statistics for Utility for *Jij Maakt Utrecht* webapplication

For the utility aspect, we regard these results as a supportive foundation for the specific requirement evaluations in section 4.2.2. Overall, *Jij Maakt Utrecht* will benefit from specific explanation section(s) and a news section. Explanations are already included, but will need to be more visible and more focussed on what exactly can be seen on the webapplication. From the government side, the wish to be able to publish news on the webapplication existed halfway through the trial. During the trial we already developed the possibility for a front-page banner. This banner has been exploited already when linking to the *Jij Maakt Utrecht* Trophy festival and the request for citizen input for further development on *Jij Maakt Utrecht*. One of the most crucial utility findings is the ability to see what is new on the webapplication. At current state, the webapplications' content is experienced as a bit stale because changes are not highlighted in one way or another. Possible improvements are the introduction of facet browsing and time-sliders options.

Another important feature that is missing is the possibility to deep-link to initiatives. Currently, the webapplication isn't suitable to create a link to a specific initiative. This is for both citizens and municipality an important requirement for upcoming period. We describe this feature in section 4.2.2 in more detail.

4.2.1.3   Memorability

On the aspect of memorability, two different sub-aspects have been measured: organisation of information and the results of user-actions. In Table 20 the results are presented.

Question topic	Mean	Median	Standard Deviation
Memorability overall	3.31	3	1.04
Clear organization of information	3.16	3	1.11
Results are conform expectations	3.47	3.5	0.95

Table 20: Descriptive statistics for Memorability for *Jij Maakt Utrecht* webapplication

These figures give us more or less a same idea about the application as what has been assessed on Learnability: the application itself is overall experienced as logically consistent. We cannot completely deduct on what level the user-experiences on this aspect are influencing the user-actions. The results become multi-interpretable when considering the organisation of information. With the information on the learnability and the feedback from our different meetings with the municipality however we do feel quite confident in the user-action logics. On the other hand, the organisation of the information on *Jij Maakt Utrecht* gives more reason for ambiguity. In our next steps for *Jij Maakt Utrecht*, we do need a strong focus on what is actually presented and how clear this actually is for citizen users.

4.2.1.4   Satisfaction

A sort of overall impression could be seen as the aspect of satisfaction. In Table 21 the results are presented.

Question topic	Mean	Median	Standard Deviation
Satisfaction overall	3.35	3	0.98
Design & layout	3.37	3.5	1.10
User-friendliness	3.41	4	0.90
Recommend to friend	3.22	3	1.01

Table 21: Descriptive statistics for Satisfaction for *Jij Maakt Utrecht* webapplication

On the aspect of satisfaction, the results are following a normal-distribution, although leaning towards a positive experience. The results of the questionnaires show us that we score a quite positive experience of the design and the user-friendliness. We need to do further research on the responses that are on the negative side. However, a couple of things can be said on these results. For one, design and layout is most likely to be subject of a wide range of subjective experiences. In software development, it is simply not possible to score positive without applying a strong design. The other side of this is that a strong design, will bring forth 'strong reactions'; positive but also negative. For us, the more interesting experience is the more passively neutral. After a more detailed look into the responses, we came on an interesting insight: the neutral responses were dominantly given by our highest age group of 50-64 years old. This over-representation of neutral experience on layout for this group biases the results in a way that the results are less outspoken positive or negative about the layout and design. Another interesting result is the difference between males and females regarding the design: males are on average half a point more positive than females. Formulated in a percentage of the maximum score of 100%, this means that males on average value the layout as 65% and females as 55%.

4.2.1.5   Efficiency

Efficiency tells us how much time after getting used to the application, is needed to perform tasks. In our surveys we asked for the user-experience on two aspects of efficiency: the consistency of the application and the straightforwardness of user actions. In Table 22 and the results are presented.

Question topic	Mean	Median	Standard Deviation
Efficiency overall	3.42	3	0.75
Consistency	3.44	3	0.56
Straightforwardness	3.41	3.5	0.91

Table 22: Descriptive statistics for Efficiency for *Jij Maakt Utrecht* webapplication

One of the most important findings of the evaluation is that new content on *Jij Maakt Utrecht* isn't prominently emphasised. The user experience on efficiency is judged on this aspect as it takes time and effort to explore the content of the webapplication each visit to check whether or not something new is published on *Jij Maakt Utrecht*. For further development, we will explore possibilities of making clear what content is new. Possible solution we are debating at the moment is facet browsing and time-slider options to be able to select the presented data.

4.2.1.6 Feedback & Errors

For usability, we also evaluate the aspect of Feedback & Errors. By this, we mean the feedback the system provides to the user in terms of actions that are available or incorrect and the options it serves for correcting user-errors. In Table 23 the results of the survey are presented.

Question topic	Mean	Median	Standard Deviation
Feedback & Errors overall	3.17	3	0.49
Clarity of feedback	3.22	3	0.55
Easy to recover mistakes	3.13	3	0.42

Table 23: Descriptive statistics for Feedback & Errors for *Jij Maakt Utrecht* webapplication

As can be seen from the table, users are rather neutral about this topic. Generally, at the on-site demonstrations of the webapplication, people were mostly engaged with discovering the possibilities and data published. They experienced no major problems and no explicit reactions about feedback & errors have been recorded. In some cases, it was not clear to people that they had to activate their account before they were able to log in and add their initiative to *Jij Maakt Utrecht*. In some cases, the activation e-mail ended up in their SPAM-folder, in other cases the e-mail has not been received at all, due to type-errors in the e-mail address. These reasons caused some confusion about the register process. Other than that, it seems people experience proper user feedback as baseline, and no strong positive or negative experienced were recorded.

4.2.1.7 Reliability

The stability or reliability of the system as experienced by the users is presented in Table 24. We asked the question whether the user experienced to be successful in the use of the system.

Question topic	Mean	Median	Standard Deviation
Reliability overall	3.22	3	0.55

Table 24: Descriptive statistics for Reliability for *Jij Maakt Utrecht* webapplication

The most important feedback we have for upcoming period is the support of tablet devices. At current state, *Jij Maakt Utrecht* is not fully supported by tablets. The reliability and the correctness of functionality on different platforms is not fully tailored for mobile devices. In upcoming period, we will investigate the possibilities for tablet support. Smartphone support is slightly different as the screens are questionably small to offer overviews on information.

Also, some users mentioned that after logging in on the webapplication, their sessions expired unexpectedly early and experienced a hindered reliability. Session-length will be looked into during the upcoming period, but early research shows that expanding the length is not ideal from a security point of view.

Apart from our survey, we explored our IIS webserver logs and come to the following findings. The distribution of the errors is as follows:

Error type	Hits
403 Forbidden	15,641
404 Not Found	817
401 Unauthorized	260
416 Requested Range Not Satifieable	47
406 Not Acceptable	24
500 Internal Server Error	1

Table 25: Distribution of errors in IIS webserver logs Jij Maakt Utrecht

We investigated these logs and from a technical point of view, there aren't critical reliability issues. Of all the errors, the majority of these errors are completely non-problematic because of auto-forwarding. Also around 8,000 (un-successful) hacking attempts are registered and did not interfere with our users' experience.

However, around 60 errors were inconsistencies in the webapplication and are considered as real errors. Also, in the total trial period, 1 internal server error has been logged, for which at the moment we have no explanation.

4.2.2   Evaluation of the requirements

After the overall user-experience evaluation on usability of the application in 4.2.1, we evaluate each of the functional requirements for the first cycle of the prototype (see also section 4.1.2). For each implemented requirement for the first trial, we present the evaluation results.

L+G-FR.03	Be able to show information in a context aware manner (location, person, time-specific, situational) to create greater awareness and create mutual citizen-government understanding

The information shown on *Jij Maakt Utrecht* is described by employees of the municipality of Utrecht as the most complete place of information. Before *Jij Maakt Utrecht*, information on co-creation spots wasn't available. In the trial, information has been collected and together with existing information been published. The map-based presentation is seen as useful by both citizens and authorities. However, the different types of information and the explanation of the different type and theme layers, is not always completely clear to users. Some information is there for increasing transparency, other information is there to increase participation. The gallery view is regarded as increasing enthusiasm for co-maintenance: the visual aspect of the current examples is encouraging, and lowers the threshold for citizens to start initiatives and share their experiences. The context information of the initiatives (labels, descriptions, initiative-owner, location, photographs) successfully met the requirement. Suggested improvements are the linkage-possibilities with Facebook, as references to Facebook have already been made in the free text descriptions of initiatives.

An important aspect of the requirements is that *Jij Maakt Utrecht* is not fully compatible with mobile devices. Users are interested in exploring information on their desktops as well as their tablets. In current state, the webapplication isn't fully supported on tablets.

Also, the possibility of a satellite map-view has been mentioned as an option to be able to see detailed information on the exact location of initiatives.

Overall requirement evaluation and refinement
L+G-FR.03 is met successfully. Next steps of investigation and development need to focus on inclusion of mobile device support (tablets in particular). Furthermore, *Jij Maakt Utrecht* can benefit from clear explanations of the type of information that is presented. A satellite option for the map-view will be considered for future development. Lastly, concrete linkage to social networks like Facebook needs to be considered.

Table 26: Evaluation of L+G-FR.03 in *Jij Maakt Utrecht*

L+G-FR.04	Provide possibilities to group, aggregate, filter large amounts of data to discover knowledge
The aim is to share and discover knowledge. *Jij Maakt Utrecht* presents large amounts of data. The key question is: what is relevant to present? *Jij Maakt Utrecht* offers filter and grouping options to users to be able to explore the information. These options are rated positively by users and the implementation is regarded as visually attractive as well, however the option to filter on time is missed.	
The webapplication is technically able to handle a lot of data, which is one of the biggest challenges for web visualisations. Conform the project planning; the real pattern recognition aggregations are not implemented yet. Data-mining functionality and knowledge discovery is planned in WP1 for the upcoming period.	
Overall requirement evaluation and refinement	
L+G-FR.04 is rated positively by the users. The webapplication is technically able to handle a large amount of data, which is one of the biggest challenges. The possibilities for the discovery of knowledge are currently on a basic level, but further advancements on data-mining and knowledge discovery are planned for upcoming period in WP1 (see the link with L+G-FR.06).	

Table 27: Evaluation of L+G-FR.04 in *Jij Maakt Utrecht*

L+G-FR.05	Provide communicative tools for government and citizens (including reactions, alerts, messages)

Overall consensus is that the requirement still stands and that it is a very important functionality of the application. Especially communication between citizen and government is seen as important; both by government and by citizens.

Usage of the possibilities to react on initiatives is regarded as low by the municipality of Utrecht as well as the citizen owners of initiatives. In meetings, it also became clear that not everyone was aware of the functionality. This gives us insightful feedback on the implementation of the requirement for further development. Also further enrichment possibilities on the aspect of communication have been derived from the evaluation sessions. Specific option to interact with Twitter and Facebook has been mentioned multiple times by both citizens and officials. Another important improvement is the possibility to share a URL to a specific initiative with others (so called "deep-link").

Regarded as shortcoming, is that currently a user has to actively look on the application to check whether or not there are reactions. Citizens and municipality both provided us with the evaluation insight that the communication would benefit from active notifications on new reactions, for example in form of an email or being able to see which reaction or initiative is new since last visit.

In the original scope of the project, it has not been assessed whether the communication should be on *Jij Maakt Utrecht*, or that *Jij Maakt Utrecht* functions more like the channel for communication. This ambiguity still exists, and continues to be a topic for internal discussions about the exact goals for the municipality of Utrecht with the application. This by itself is of course a desirable effect of the co-creation process: agenda-setting of important topics.

Overall requirement evaluation and refinement
The implementation of the requirement L+G-FR.05 as assessed initially is experienced positively. However, during the co-creation process and trial, new ideas and new insights on the aims for the requirement provide us with possible enrichments and specifications on the aspect of communicative tools. Concrete deep-link options need to be investigated for future development as well as forms of active notifications for new reactions and initiatives.

Table 28: Evaluation of L+G-FR.05 in *Jij Maakt Utrecht*

L+G-FR.06	Provide possibilities to visualize knowledge and patterns from data by data-mining techniques

The requirement is essentially met successfully. However, this is taking in regard what is possible technically. As described for L+G-FR.04, the possibilities to group and filter large amounts of data are met positively. The way the results of filtering and clustering are presented is regarded as visually attractive. However, The actual knowledge and patterns that can be visualised are on a rather basic level.

With upcoming work in WP1 on data-mining and knowledge discovery, we are confident that with our Webapplication for eGovernment Dialogue and Visualisation, we have a solid foundation in the current webapplication. The true innovative character will need to prove itself in upcoming period, based on the possibilities to recognize interesting patterns.

Overall requirement evaluation and refinement
The requirement is met successfully on the aspect of the visualisation. The content and patterns presented are currently not state-of-the-art yet; however this conforms to the Live+Gov planning. Further developments in upcoming period will create the possibilities to present real knowledge patterns. L+G-FR.06 is currently at a stage where it proves itself as a firm foundation for future developments.

Table 29: Evaluation of L+G-FR.06 in *Jij Maakt Utrecht*

L+G-FR.13   Be able to connect to different municipality administration systems
It was planned to have direct input in Jij Maakt Utrecht with administrative data. However, the data in mind (by municipality) was simply not available in any type of format. Thus, a connection to the administrative system wasn't possible. In the trial, we have done data-collection and manually imported this data into the webapplication to overcome the issue. Functionally, this had the same result as a connection. From the integration point of view, we have to conclude that the requirement was not met. Furthermore, as one of the important insights we gained while the co-creation process progressed is that contrary to initial information, the availability of data is minimal. This makes us question the goal of the requirement. For the moment, L+G-FR.13 has low priority for the *Jij Maakt Utrecht* scenario.
Overall requirement evaluation and refinement
Interesting insights have been gained from the trial. Contrary to initial assessment, data interesting for Jij Maakt Utrecht is not available in any format that fits a true external connection. For the trial, we overcame this by organising data-collection sessions with the customised data-collection prototype. However, at this point we assess L+G-FR.13 as having low priority for *Jij Maakt Utrecht*. This means that we discontinue L+G-FR.13 for Jij Maakt Utrecht, because of practical reasons. We do want to note here that we are confident that our architecture is perfectly able to connect to external administrative systems.

Table 30: Evaluation of L+G-FR.13 in *Jij Maakt Utrecht*

L+G-FR.20   Provide possibilities to present co-maintenance initiatives in an area
This of course is the core requirement for the first trial. The feedback on the presentation of the co-maintenance spots is very positive. From the citizens' perspective, the overall reaction is that it is really nice to see what is actually going on in their neighbourhood. People are surprised with the amount of co-maintenance and the way of displaying this information is met with praise. Especially the combination of the gallery view and the map-based view is applauded. For the authorities, the requirement has been a starting point to collect the actual information. *Jij Maakt Utrecht* is the first place that has the actual information about co-maintenance and it is also described as a complete source for information. The requirement has been fully met. There is a clear link with L+G-FR.25.
Overall requirement evaluation and refinement
L+G-FR.20 has been fully met. Both citizens and authorities praise the completeness of information that wasn't available at all before *Jij Maakt Utrecht*. The combination of the gallery view and the map-based view is regarded as both informative as well as visually attractive.

Table 31: Evaluation of L+G-FR.20 in *Jij Maakt Utrecht*

L+G-FR.25	Improve the quality of existing (open) data sets related to urban space, by leverage of citizens functioning as sensors and verify with existing registrations

Important insight for this requirement has been that the presumed existing data regarding co-maintenance initiatives was actually very limited to non-existent. This became clear when the application was ready to be provided with the content, which the municipality committed to provide. After multiple sessions, it became clear that the data could not be provided. Specifically for this issue, namely that there was no content to show, the customised data-collection prototype has been developed. With this mobile prototype, we had specific data-collection trial in which the co-maintenance spots of Utrecht have been registered. 420 spots have been registered including photographs. This data collected with the mobile application has been published on the webapplication. This way, the leverage of citizens as sensors has been successfully tested: spots that were not yet registered, are now not only registered, but also published to public.

Also, other data-sets have been published on *Jij Maakt Utrecht*. For these sets, the citizen as sensor approach hasn't been applied and this is also not in our current scope. The data sets have been downloaded more than 600 times in total. During the trial, new sets have been provided by Utrecht and were published on *Jij Maakt Utrecht*. This shows that *Jij Maakt Utrecht* plays a role for the municipality of Utrecht in regards of their data sets.

Overall requirement evaluation and refinement

For the scope of the Urban Maintenance trial, *Jij Maakt Utrecht* is a catalyser for collecting data about co-maintenance. New insight for the municipality was that this data was not available. We can conclude that we met L+G-FR.25 beyond success as we gathered the data with the citizen-as-sensor approach and not just improved the quality of existing data. Apart from data within our scope, *Jij Maakt Utrecht* holds a diverse range of data-sets. The ambition to improve the quality of all these sets is too high. On the level of transparency, the publication is successful. However, the municipality of Utrecht concludes that we should be careful with bringing too broad goals into the application. As Use Case owner, we agree with the strategy that it is better not to expand to too many different domains within the webapplication.

Table 32: Evaluation of L+G-FR.25 in *Jij Maakt Utrecht*

L+G-FR.26	Introduce elements of serious gaming / gamification

Serious gaming has been mentioned at the start of the co-creation process of *Jij Maakt Utrecht*. At that point, no concrete ways to implement gamification aspects in *Jij Maakt Utrecht* were assessed. The gamification concept was met with enthusiasm of the municipality of Utrecht. Along the process, gamification has been discussed and explored. An experiment has been set up and executed. Although the experiment has been met with enthusiasm of stakeholders and media, the concept of gamification has been assessed as too ambitious by the municipality of Utrecht.

Overall requirement evaluation and refinement

The concept of gamification has been explored and an experiment has been executed. Although perceived as interesting, the requirement is regarded as too ambitious by the municipality of Utrecht. This means that we discontinue L+G-FR.26 in Live+Gov.

Table 33: Evaluation of L+G-FR.26 in *Jij Maakt Utrecht*

4.2.3 Evaluation results for specific WPs

4.2.3.1 WP1 – Reality Sensing and Mining

At current state, WP1 has no explicit role in the Urban Maintenance Use Case. However, future development of data-mining techniques will be real important. Advancements in personalised content delivery and techniques for knowledge discovery will provide the project with possibilities on higher context extraction. From here, we will be able to visualise knowledge patterns that apply to our L+G-FR.06. In current state, our trial users are positive on the visualisations, but regard the level of information as being in a rather basic state. From our Use Case, we can only re-state the need of the planned WP1 work for upcoming period. With the starting level of the patterns we are able to show already, we prove that the foundation is firm and we are eagerly anticipating to benefit from the upcoming work of WP1.

4.2.3.2 WP2 – Policy Modelling and Visualization for Mobile-aware Dialogue

The entire process involved with *Jij Maakt Utrecht* is highly important for WP2 and future training package material. The organisational and political embedding of the webapplication is not a straightforward process. With Live+Gov we are active in the transitioning towards a we-government approach. Utrecht is regarded as front-runner on the professionalising of participation initiatives, but even here a lot is not fully clear. Within Utrecht there is an entire process that revolves around the question how to facilitate participation and what is the role of the municipality in the new we-government approach. *Jij Maakt Utrecht* plays an important role in this transitional re-invention. During the trial, a lot of documentation and information on the organisational level of embedding participation in the administration has been gathered. Very important is the insight that we are in the middle of the transition. Currently, we are in the middle of internal political and organisational self-realisation along with the new participatory demands of citizens and government itself. With our Use Case, we are able to provide WP2 with the feedback and information regarding the transition and shifting roles from the actual real life-world. This information has been provided to WP2, which included and theorised on this material in D2.3 [12].

Note has to be made that gathering this type of information is very resource intensive. As can been seen from our information gathering moments, a lot of time has been invested on co-creation (see Table 16). The processing of all this rather unstructured information has shown itself as a very intensive task.

Apart from the more conceptual organisational and political input for WP2, the evaluation results are mainly related to *Live+Gov Milestone 12: Visualisation Tool for eParticipation for first trials*. These results are too extensive to present at this section: the usability results for the *Jij Maakt Utrecht* are of great importance for *Live+Gov Milestone 15: Visualisation tool for eParticipation for second trials*. See sections 4.2.1 and 4.2.2 for more detail. Overall, the *Jij Maakt Utrecht* provides WP2 with user experience on the tool and provide in depth information for the WP2 Training package and implementations of eParticipation scenarios on an organisational level.

4.2.3.3 WP3 – Mobile Augmented Reality

For this first field trial, we have consulted our WP partners to assess what information would be beneficial for WP3. In the first field trial of Urban Maintenance no augmented reality components are being used in the prototype and no concrete evaluation information needs

were defined. Subsequently, we do not have concrete evaluation results specifically for WP3.

4.2.3.4   WP4 – System Integration

Integration

The JMU solution strongly depends on account management and access control mechanisms. These mechanisms are applied to distinguish between functionality for the public and functionality for government officials. These SaaS Services Center modules proved to be very effective, allowing well-structured account management and access control for both citizens and government officials, without raising invisible barriers for participation. During the trial, one of the findings is that one or more extra can and should be defined and configured for the next trial.

Scalability / availability

JMU is launched as a public beta, which requires very high demands on stability and availability. Extra attention is given on these aspects during the design and verification of the eGoverment Dialogue and Visualisation Website. The followed technical approach (Single Page Application) proved to be very effective. Initially the media attention during the launch caused peaks in the visitor numbers, but no noticeable performance degradation was experienced. Section 4.2.4.1 gives more insight in the exact user and usage numbers.

(Mobile) Browser compatibility

Another import aspect in public trials is cross-browser compatibility. In preparation of the launch extra attention is given to these aspects in the test procedures. The test verified that not all HTML5 functions are supported (e.g. data validation) properly by all browsers yet. Browser specific alterations were required to support the popular browsers (Chrome, Safari, IE, Firefox).

Although HTML5 support is increasing [20], this will be an important attention point for public web applications created with the Live+Gov Toolkit.

For the first trial JMU is primarily targeted on desktop usage, however future mobile usage is envisioned. The tests show that in its current form the application experiences some hick-ups in some mobile browsers. For the 2nd trial extra attention will be given to mobile browser compatibility.

Big-data optimization

JMU has two main presentation options: 1) the map screen and 2) the photo gallery view. Both views are designed to show large amounts of data, using mechanisms like client-side aggregation, which proved to be very effective during the trial. The foreseen growing amount of data will further raise the performance requirements for these components. Additional test performed with larger amount of data show that further optimization is required. Techniques like server-side clustering and lazy-loading need to be investigated to optimize the performance when working with large data sets. [21]

Data sets

During the preparation of the trial the options were investigated to connect to the public administration system to import the dataset with co-maintained spots. Technically the

connection options are present, however the content itself proved to be more challenging. Spots maintained by the government are well inventoried and centrally stored, however the administration of co-maintained spots is fragmented and often incomplete. Before starting the trials a customized version of the issue-reporting client was developed to gather the required data in the field together with the public officials.

Analytics

JMU is built as a single page application, which has advantages for scalability, but also causes some inherent barriers for enabling analytic tools like Google Analytics. For these tools the JMU is seen as one (single) page, therefore no additional data is available about the usage of the individual functions and pages inside JMU. For the 2nd trial the options for more advanced technical analytic options need to be investigated.

4.2.4 Impact and success of the Use Case trial

4.2.4.1 User and usage numbers

Statistics about visiting behaviour at *Jij Maakt Utrecht* are derived directly from the database.

Indicator	Number
Number of unique visitors	4,777
of which in the Netherlands	2,703
of which from IP municipality Utrecht	601

Table 34: Number of visitors at *Jij Maakt Utrecht* webapplication

From the report, it can be seen that about 60 percent of the unique visitors are situated in the Netherlands. So, a total of over 2,700 unique visitors are from the Netherlands at *Jij Maakt Utrecht*. During the trial period, some peaks are visible in the daily visits (see Figure 11).
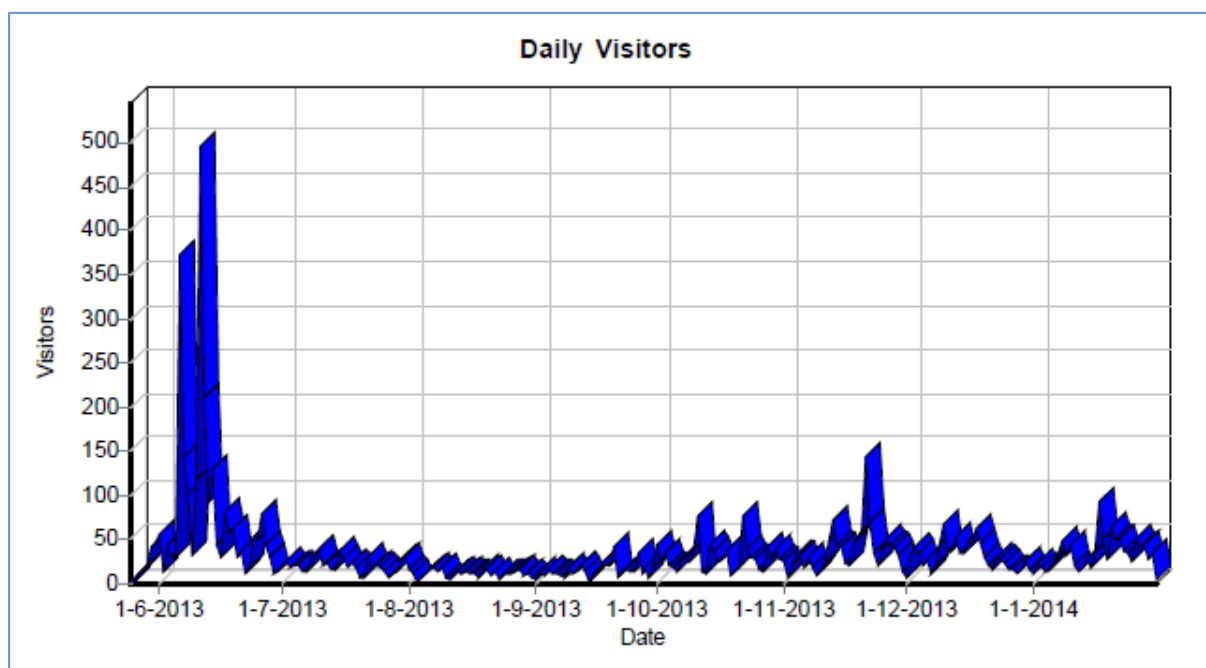
Figure 11: Graph of daily visitors of *Jij Maakt Utrecht*

After the launch, two large peaks are visible of 400-500 visits in one day. The launch was at May 30, 2013 and the extended launch at June 1, 2013. It can be clearly seen from the graph that the two peaks are after these dates. In the summer period, visits were very low. On November 21, 2013, the *Jij Maakt Utrecht*-festival took place in Utrecht, where a trophy was awarded to a citizen initiative. Around this festival, an increase in the number of visits can be clearly seen up to about 150 visits in one day.

In the first six weeks, *Jij Maakt Utrecht* was accessed many times. It can be clearly seen that the most visits were made in the week of June 9 to 15, 2013 (see Table 35). This was 1-2 weeks after the launch.

Week	Hits	Page Views	Visitors
26-5-2013 – 1-6-2013	58,472	12,873	191
2-6-2013 – 8-6-2013	56,191	20,650	739
9-6-2013 – 15-6-2013	95,431	35,502	1,123
16-6-2013 – 22-6-2013	25,078	8,307	309
23-6-2013 – 29-6-2013	24,599	8,078	287
30-6-2013 - 6-7-2013	10,623	3,279	131

Table 35: Activity per week in the first six weeks of the field trial

Reactions

In *Jij Maakt Utrecht* it is also possible to react on initiatives. To keep this easily accessible it is not necessary to register first if you want to give a reaction. A total of 42 reactions to 19 different initiatives have been given in the trial period.

Users

In total, 120 users are registered at *Jij Maakt Utrecht*. When users register, they need to confirm the registration by e-mail to activate their account. With 8 users this did not happen; this means a total of 112 active registered users at the webapplication. Registration is necessary when someone wants to add an initiative to *Jij Maakt Utrecht*. For visiting the application and/or placing a reaction, it is not necessary to register (see also D5.2 [1]). Therefore, the number of users should not be confused with the number of visitors presented above.

Initiatives

Only active users are allowed to add an initiative to *Jij Maakt Utrecht*. 56 active users have added 1 or more initiatives to the webapplication (see Table 36).

Number of initiatives per owner	Number of owners
1	47
2	5
3	2
4	2
Grand Total	56

Table 36: Number of owners that own 1, 2, 3 or 4 initiatives at *Jij Maakt Utrecht*

In Figure 12, the distribution of the different type of initiatives is presented. These numbers include citizen and public officials.
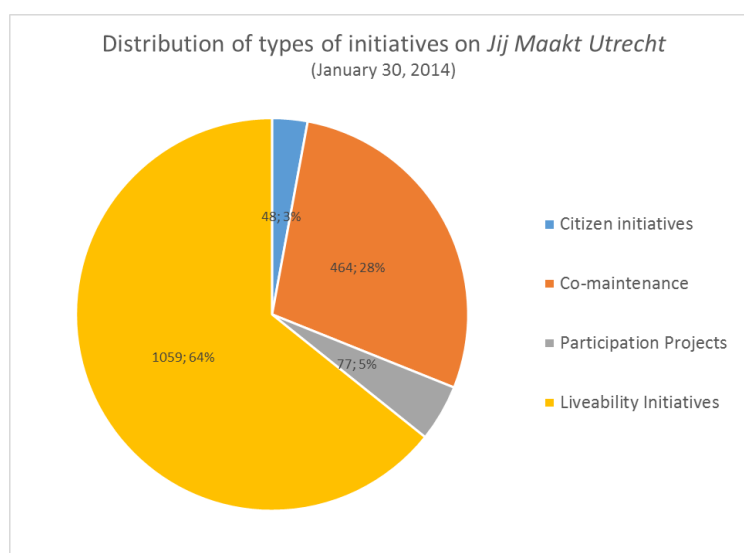
Figure 12: Distribution of types of added initiatives for *Jij Maakt Utrecht*

For our trial period, we see quite an equal balance between the initiatives placed by citizens and those placed by public officials (see Figure 13).
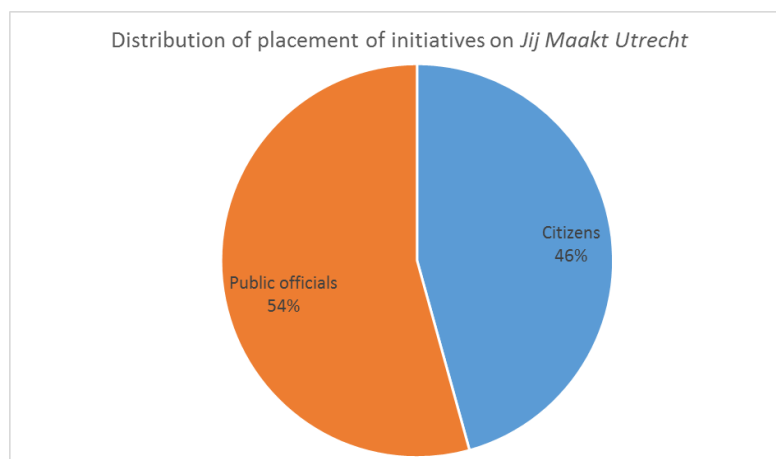
Figure 13: Distribution of initiatives added by public officials and added by citizens

Downloads of open data sets

The numbers of downloads of published open data set is shown in Table 37.

Open data sets	Number of downloads[2]
Street furniture	204
Boundaries	186
Street names	158

Table 37: Downloads of open data sets at *Jij Maakt Utrecht*

Participation festival: *Jij Maakt Utrecht* festival

The impact of *Jij Maakt Utrecht* should be measured not only based on numbers but also on the embedding in the city administration. With the philosophy to take an evolutionary approach, the actual 'integration' of *Jij Maakt Utrecht* in the city administration is an important aspect of the entire trial. There is a relation to most recent deliverable of WP2 (*D2.3 - Visualization of data injection from mobile sensing*), in which it is described that the visualisation of participatory processes is bringing structure to informal structures of decision-making.[12] For the municipality of Utrecht, this is quite a successful showcase.

One major sign of this success is the role *Jij Maakt Utrecht* plays in the overall participation policies for the municipality. The webapplication is a central topic on the overall internal vision-shaping process that is taking place in Utrecht. This has been extensively described in the previously mentioned deliverable D2.3. As a concrete sign of success of impact is the renaming of the yearly "Participation festival". This festival has actually been renamed to the *Jij Maakt Utrecht* festival, a clear sign of our solution taking a significant role in the policies of the municipality of Utrecht. With the name change, also a slight change of approach took place: where the festival used to be mainly an internal affair, this year the festivals focus is on the citizen itself. This clearly shows the internal perspective of the municipality becoming broader and more citizen-orientated: a true achievement.

---

[2] Based on number of times the button 'download' was pressed.

European impact

Last but not least, *Jij Maakt Utrecht* has been picked up on a European level as well. In November 2013 the document 'Cities Supporting eInclusion and Citizen Participation' by EUROCITIES got published. EUROCITIES is a network of major European cities which brings together 130 of Europe's largest cities and 40 partner cities that aims to reinforce the role that local government should play on day-to-day citizen level. The report explores how improving technology can benefit the lives of citizens. In the report the focus is on participation and the report gives outlines and examples of cases of successful initiatives all over Europe. The publication reports on 8 successful showcases of which *Jij Maakt Utrecht* is one. Quotes taken from this leading report [19]:

- Jij maakt Utrecht' ('You make Utrecht') is part of Participation 2.0, an innovation platform that focuses on social experimentation in the city of Utrecht. The platform uses new digital technology to support local democracy, encourage active citizenship and support people to self-organise digitally. This initiative is part of strengthening and broadening citizen participation in Utrecht.

- Successes - The first statistics on the use of the website were collected just three weeks after it was launched. During this time, the site attracted more than 1,500 visitors. Six articles appeared in various media outlets and over 40 tweets were generated. Almost 20 people posted their feedback on the initiative.
    - The first feedback is very promising. Citizens are engaged, people are talking about it and most of the opinions and feedback received are very positive. One example of an opinion stated: "As a citizen of Utrecht I am quite proud of this. The approach of the portal – citizen participation – is also promising."
    - On the website you can find:
        - 979 liveability initiatives
        - 422 projects where residents themselves manage public space
        - 72 municipal participation projects
        - 24 initiatives from residents to create something together
        - several open data sets.

The report also saw a challenge for *Jij Maakt Utrecht*:

- Participants praised the digital approach, but also indicated the need for more discussions with the municipality both during and after the online debate.

4.3   Summary of the evaluation of *Jij Maakt Utrecht*

The *Jij Maakt Utrecht* field trial has been running in a public beta for 8 months. Moreover, the trial started even before the launch of the public beta. The preceding co-creation process with the municipality of Utrecht, including the initial data collection, also generated many important evaluation results. Overall, the first trial of *Jij Maakt Utrecht* is considered a success.

The most important evaluation result is that all stakeholders are very impressed with *Jij Maakt Utrecht* and the possibilities of the platform. Much positive feedback has been gathered and stakeholders are convinced that the webapplication is very useful in practice. Next to the mainly positive feedback, constructive feedback is also gathered. Regarding communication, an often heard critique is that it should be very clear to users what the exact purpose of *Jij Maakt Utrecht* is. This corresponds to the changing role of the municipality, which the municipality is shaping, as a part of the wider transition process to we-government.

Next to that, an important result that has been gathered during the co-creation process is that it is challenging to gather content and data to initially fill the webapplication. To tackle this point, among other efforts there has been a data collection event for co-maintenance objects. Furthermore, another point that was necessary before the launch of the public beta was that moderation should be in place on this administrative platform.

Our user experience evaluations show that *Jij Maakt Utrecht* is met with very positive reactions on usability. A mentioned improvement in usability is the support of (mobile) tablet devices. It also became clear that the dialogue platform will benefit from possibilities of deep-linking in order to provide possibilities of easy crowd-sourcing. Furthermore, both citizens and government officials experienced difficulties assessing the dynamic character of *Jij Maakt Utrecht*, i.e. what content is new since last visits. On this aspect, we see features of facet-browser and time-sliders as crucial improvements for the next period.

The trial gained noticeable attention in media and political bodies on a local, national and European level. For Utrecht itself, the process of the transition towards a more we-Government organisation has been leveraged considerably, with *Jij Maakt Utrecht* as an important platform and point of reference in this transition process. This is an important trial result: *Jij Maakt Utrecht* and the co-creation process in which it is developed, supports this process of wider transition within the municipality of Utrecht. It really helps to (re)define roles, processes and agendas within the municipality.

The preliminary analysis on data already shows that state-of-the art patterns are expected to discovered. During the upcoming period it is planned that WP1 achieves considerable progress on this aspect, which will be valuable input for web visualisations.

The municipality is constantly reflecting and taking next steps on their self-realisation in the transition. The many internal sessions that were held provided insights that the new role is not a trivial one to define, especially since it involves very diverse stakeholders. Information gathered during the trial is of important value for WP2 and Live+Gov training package. Key insight is that, next to developing tools for a defined eParticipation organisation, Live+Gov is part of the transition towards we-government. This point is important, as what we have seen is that the goals and purposes are shifting as a result of the trial.

Overall, the Urban Maintenance field trial with *Jij Maakt Utrecht* has reached a large impact. *Jij Maakt Utrecht* enlarges engagement and triggers opinions, as has been published in several press and blog articles about *Jij Maakt Utrecht*, self-organisation and participation. Moreover, an important European report stated the webapplication as key example for eParticipation initiatives and the embedding of *Jij Maakt Utrecht* in the organisation of Utrecht is strong..

Finally, it is important to note that the stability of the webapplication has been tested in public. Technically, the developed platform has shown itself as very stable throughout the entire trial. This point is of course of importance for real end-users solutions and is of great importance in regards of our SaaS-based approach. With considerable usage and data, the performance of the customised webapplication for eGovernment Dialogue and Visualisation has proven itself as firm foundation for the next steps of Live+Gov.

5 Urban Maintenance Use Case: Issue reporting and reality mining

In this section we present the evaluation of the first field trial of Issue reporting and reality mining scenario for Urban Maintenance. First, in section 5.1, we will describe the field-trial set-up (including a specific City Safari event), give a brief overview of the prototype used, and the tailored evaluation process. Following the general evaluation strategy of WP5 (section 2), we present results on the usability of our customised prototype, evaluate the requirements, describe the Use Case impact and give specific input for WP1-WP4 in section 5.2.3. In section 5.3, we conclude the Urban Maintenance Issue Reporting trial with an overall summary of the trial and the results collected.

5.1 Urban Maintenance Field Trial

5.1.1 Field Trial set-up

5.1.1.1 Extended feedback mechanism

The existing BuitenBeter application has been extended with an advanced feedback mechanism. This prototype extends the feedback-loop significantly: apart from a status update from the municipality on reports, now the municipality has the option to send textual feedback, which makes the BuitenBeter application a two-way communication application about maintenance topics in urban space. It involves a web-interface for providing feedback and the mobile application BuitenBeter, which constitute the customisation prototype delivered in D5.2. In the Issue Reporting client (which is running for the entire Netherlands) the feedback is dynamically shown in the application when provided by the municipality. In other words: it is the same production application, but based on which municipality the report is sent to, the textual feedback is available.

For the field trial, we have a selection of 13 municipalities and one specific harbour area of Rotterdam for which we activated the textual feedback mechanism. This means that for reports that are sent to these specific municipalities, the authorities are able to provide the citizen with status feedback and textual information about the (processing of the) report. One municipality (Eindhoven) has an automatic connection to their administrative system in form of a StUF-connection.[6] The trial is taking place in the following set of Dutch municipalities:

Municipality with extended feedback mechanism option	Number of inhabitants (2013-1-1) [7]
Bergen op Zoom	66,287
Boekel	10,062
Borsele	22,683
Eindhoven[3]	218,433
Etten-Leur	42,274
Gilze en Rijen	25,858
Halderberge	29,231

---

[3] BuitenBeter has a direct connector to the administrative system of Eindhoven

Heusden	43,244
Huizen	41,445
Moerdijk	36,625
Ommen	17,314
Roosendaal	77,155
Havenbedrijf Rotterdam[4]	616,294
Sittard-Geleen	94,024

Table 38: Trial municipalities with extended feedback mechanism

The trial started end of March / beginning of April 2013 and went on for the rest of 2013. The extended feedback mechanism option is still active.

As we are exploiting the existing production application that runs on Android, iPhone, Windows Mobile, Windows Phone, BlackBerry and Symbian devices, the recruitment of trial users followed a very natural path. Existing and new users got the possibility to receive extended feedback seamlessly with their application. For the 13 municipalities, we actively informed our contact persons in the administrations and provided them with the information and concepts necessary for proper usage of their possibility to provide the citizen with insightful feedback on their issue reports.

The trial is running in production environment where we have to act carefully with functional and technical innovations for obvious socio-political reasons. The upside of this is that we are able to tap into a large existing user-group on a running system that is well-known. For possibilities to take steps forward without risking production situations, we ran a specific trial for Issue Reporting: Jekerkwartier City Safari. This specific event has also been used for the qualitative side of the Issue reporting evaluation to gain insights on usability aspects of the applications. More on the specific event in section 5.1.1.2.

5.1.1.2 Jekerkwartier City Safari

We assess our route to make impact carefully; one important aspect is that in a production environment, we have to act carefully as mentioned already in section 5.1. We took the step (that would be too precarious to take in the running BuitenBeter system) to integrate the Issue Reporting application with the SaaS Service Centre and the Issue reporting web application for a specific City Safari event.

In the area of Jekerkwartier in Maastricht, BuitenBeter was used for a 1-day trial. In this area, south of the city centre of Maastricht, about 1500 inhabitants live and 120 entrepreneurs are situated. Next to that, many cultural buildings, university departments, theatres and the conservatory are situated in this area. The Foundation for Entrepreneurs in this area is determining a strategic vision about this area and therefore has high interests in getting information about how a representative group of people experience this area.

To capture experiences, the foundation organised a 1-day inspection of the area with a group of 50 people: a City Safari. This group consists of citizens, Dutch and foreign students, entrepreneurs and people from other areas in and outside of Maastricht. Before the

---

[4] Although under legislation of the municipality of Rotterdam, the extended feedback mechanism has been activated for the harbour area of Rotterdam and doesn't encompasses the entire municipality of Rotterdam

experience mining, an introduction would take place including an introduction by a historian about the neighbourhood. See for an impression for the introduction Figure 14.
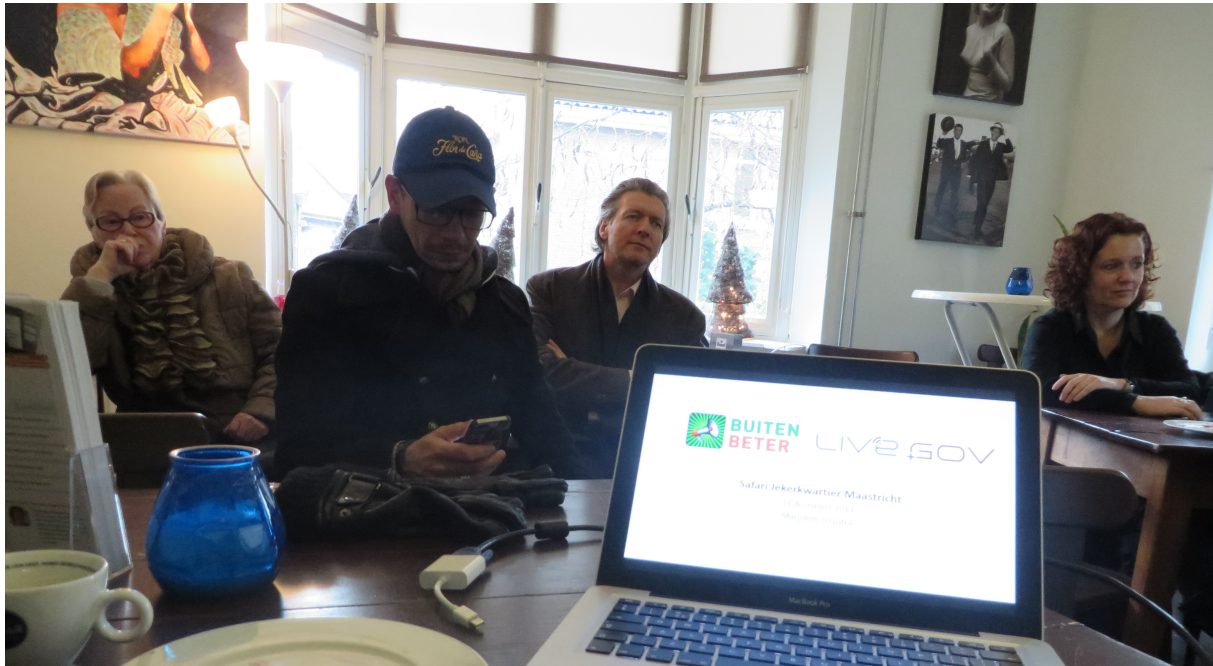
Figure 14: Introduction of City Safari

The idea for this event is that experiences at a certain location are captured in pictures, which will be rated. BuitenBeter perfectly fits these requirements: picture, location and category in a few steps. The only requirement is that the participants own a smartphone. Possible participants were invited by e-mail for the city safari. In addition to that, an event was created on Facebook.

City Safari trial

Saturday December 14, 2013, the trial took place. About 15 participants showed up, in spite of the rain. The historian introduced the area to the entire group and an instruction of the day was given. After this plenary session, the group split up in smaller groups that went on safari in the area of Jekerkwartier. During the afternoon, a total 127 reports were made on different locations in the area. The results can be seen in Figure 15 and Figure 16; the customised visualisation of the Issue Reporting data, based on the Webapplication for eGovernment Dialogue and Visualisation component (Live+Gov component C18, see for more information D4.1 [11]).
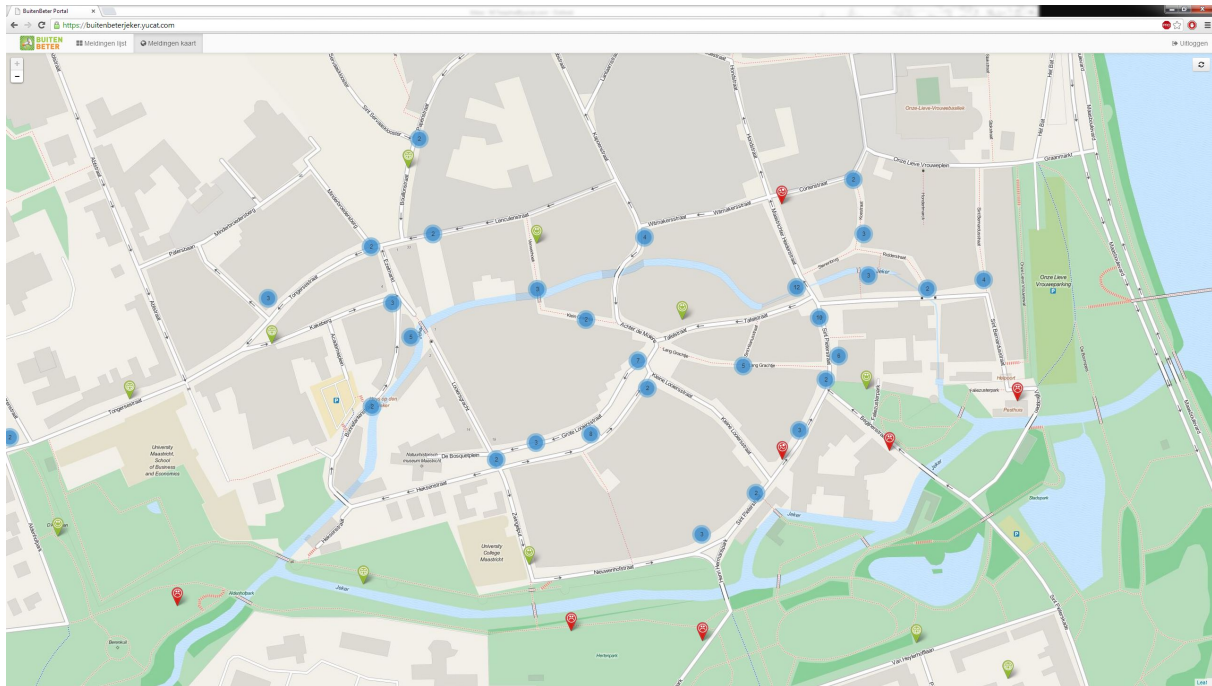
Figure 15: Overview of reports in Jekerkwartier Maastricht, December 14, 2013 (the number represents the number of reports in that cluster)

Figure 16: Customised tags for reports on the same location, three negative and two positive

The City Safari trial was held with several goals. The stakeholder groups' main interest was to receive rich input for their strategic vision for the neighbourhood. Apart from this non-governmental policy modelling information, another important need was assessed. The trial was aimed to create engagement among citizens and other stakeholders of the domain of public space. From the perspective of the Use Case owner Yucat, the trial ought to leverage the dedicated use of BuitenBeter. This goal is shared with the municipality of Maastricht, which searches for methods to increase the number of issue reports to get better insights in the citizen's perspective and the current state of the maintenance in their municipality. Moreover, the trial itself is seen by all involved parties as a good opportunity for publicity for BuitenBeter and participation in general. In the scope of Live+Gov, we see the trial as an important event to gather further insights in the Use Case user experience for our solutions. Internally, the specific customisation for the City Safari gives knowledge and confidence in

integration of the Use Case with the Live+Gov architecture as well as the customisation possibilities.

The set-up of the trial counted on 50 users. This target has not been met with the 15 participants. Although enthusiasm is high, actual active participation is not a trivial job to achieve. The number of trial users might be disappointing on the first look; the participation of a council member on the other hand, is a highlight. In general, it can be said that both citizens and other stakeholders see events like this, that increase citizen participation and experience-mining, interesting as well as important. The actual practicing on said high level goals however, is not easy to increase.

This last notion is important for the entire project. There is a gap between regarding participation as important and the actual commitment to it. This is an important duality for the domain Live+Gov is active in. On a small scale this holds for the trial participation itself, for the larger picture of Live+Gov it is needed to have less ambitious and more realistic targets or form a strategy to increase participation numbers. Most likely, the solution will be a combination of the two: the experiences gathered from the trials will make more precise estimations possible and for a stronger embedding of participatory initiatives in society WP2 will help to form strategies and best practices.

5.1.2 Issue reporting prototype

The prototype used for the first field trial has extensively described in D5.2. In this section we give a short overview of the prototype. For the implementation of the prototype the following requirements out of the complete Use Case requirements, as assessed in D5.1, were defined as crucial for the first pilot [1] [2]:

#	Live+Gov functional requirement
L+G-FR.02	Be able to sense the context of users (e.g. location, means of transportation, activities etc)
L+G-FR.05	Provide communicative tools for government and citizens (including reactions, alerts, messages)
L+G-FR.08	Provide possibilities for textual feedback. For example how his contribution is taken into consideration, including possible benefits for participating and status report on the situation
L+G-FR.09	Allow a user to view all of the issues that are currently open, and also a history record of "my issues" (where the citizen has participated) and "my municipality" (where all of the past issues from the town can be found) with a short summary of each of the past issues.
L+G-FR.10	Be able to utilize 'citizens as sensors' for their environment by mining relevant data (e.g. location, context, pictures, reports) from their reality
L+G-FR.13	Be able to connect to different municipality administration systems
L+G-FR.19	Collect and mine citizen experience and citizen (implicit) norms

Table 39: Requirements for first Issue Reporting trial

In Table 40 the involved Live+Gov components for the first Issue Reporting trial are summarised [1]:

#	Live+Gov component	Type	Description
C1	Issue Reporting Service	Back-end web services	Issue Reporting (BuitenBeter) Service. Back-end Web Service for the mobile data-collection prototype, which allows filing a report.
C2	Issue Distribution Service	Back-end web services	The server side service for distributing the specific data collected with the mobile data-collection prototype. For BuitenBeter the distribution service handles the specific allocation of the correct municipality of the reported issue.
C3	Issue Updating Service	Back-end web services	Issue Reporting Updating Service. Updates issue report status information. For the extended feedback prototype, this service has been advanced with the possibility of textual feedback.
C11	Mobile Issue Reporting Client	Mobile Application	BuitenBeter extended with the possibility to receive textual feedback of municipalities on the specific issue.
C16	Webapplication for Enhanced Issue Report Feedback	Webapplication	Webapplication for enhanced issue report feedback enables a 'low-tech' municipality to give feedback to issue reports. They receive an e-mail for the issue reporting distribution service, including a link to this webapplication. In this webapplication one can change the status of the issue report and with the extended feedback mechanism which is prototyped for the Urban Maintenance Use Case; also a free text message can be added for the citizen that submitted the report.

Table 40: Live+Gov components for the first Issue Reporting trial

Prototypes used for the City Safari

The 1-day event in a restricted area opened up possibilities not available in the production environment. For the trial, we did specific customisation on top of the BuitenBeter production client:

Customisation for City Safari prototype
A new municipality added: Maastricht Jekerkwartier
4 specific tags were customised, dynamically available for the specific area: Super Top, Top, Flop and Super Flop (ratings of the experience from very positive to very negative)   These tags are *dynamically available* based on the location of the citizen for the specific area.
Visualisation of the mobile issue reports in the Issue reporting web application   • A map based view   • An image based view
Integration with the SaaS Service Centre
Specific role for the municipality of Jekerkwartier in the issue reporting web application.

Table 41: Issue reporting and reality mining prototypes for City Safari

5.1.3 Evaluation process

Results from the 9-month period are used for our thorough evaluation, which is mainly a quantitative analysis and, on the contrary to the evaluation of *Jij Maakt Utrecht* (see section 4), it contains less qualitative research. Focus of the evaluation for the Issue Reporting trial has been the impact in numbers and the embedding of the extended feedback mechanism for BuitenBeter in citizen- and administrative usage. For the highly quantitative evaluation approach the main source of data for the evaluation is the BuitenBeter database.

We have conducted a full data analysis of more than 500,000 database records for Issue Reports. With the massive amount of new and existing data, we are able to analyse and compare in a convincing way. That said, data analysis by itself is not always a straight route to knowledge discovery. In most cases, results are multi-interpretable, include biases and ask for assumptions. For transparency reasons, we provide the results and examples first and try to form interpretations and conclusions separately. It should be noted that the explanations for the statistical results are based on more than just the figures as we have a long relationship with the domain of urban maintenance. It is not always possible to acknowledge these more informal and experience based sources.

Apart from analysis on the usage of the features, we have done surveys for specific usability measurement, for which we exploited the Jekerkwartier City Safari event (see 5.1.1.2). The surveys have been set-up based on a tailored combination of 9 different usability questionnaire survey standards (see Table 1) that we've analysed. With our surveys, we aim to get user feedback for our prototypes on the different usability aspects of Table 42, as set out in the Evaluation Strategy (section 2) as well.

Aspect	Description
Learnability	How easy is it for users to accomplish basic tasks the first time they encounter the design?
Utility	Does it do what users need?
Memorability	When users return to the design after a period of not using it, how easily can they re-establish proficiency?

Satisfaction	How pleasant is it to use the system in general?
Efficiency	Once users have learned the design, how quickly can they perform tasks?
Feedback & Errors	How many errors do users make, how severe are these errors, and how easily can they recover from the errors?
Reliability	Does the software runs consistently without crashing?

Table 42: Usability aspects

The survey consists of 33 questions, thus each aspect is covered by multiple questions. The surveys have been set-up in Dutch and make use of a bipolar Likert-scale [10]; measuring positive or negative responses to statements:

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

The answers on multiple questions are grouped and aggregated to give convincing insights on each topic. These topics combined give a rich idea about the overall usability of the application.

In sum, for the evaluation of the first Issue Reporting trial, we undertook the following evaluation activities:

- Analysis of data of more than 500,000 database records of Issue Reporting.
- A survey questionnaire distributed to 15 people with specific usability and user experience questions for City Safari. 7 surveys have been fully completed and returned. Apart from this research, the application BuitenBeter is running successfully for years and we have a lot of different iterations sessions behind us already based on user experiences. In other words, the usability aspect of the Issue reporting client has been tested and evaluated by a large audience for a long period. This feedback has been used extensively for adjustments and improvements.
- Analysis of appstore reviews, Twitter and emails of user experience feedback

The following Table 45 summarises the evaluation aspects for this use case on a high-level. For each aspect the used evaluation methods and the WP the evaluation of the specific aspect is delivering results to is given.

Protoype evaluation				
Aspect	**Criteria**	**Specifications**	**Targeted WP**	**Methods used**
Customization process success	Integration	How well were included components integrated?	WP4 & WP5	Internal review

	Compatibility	How well does the application perform on different mobile platforms	WP4	Internal review
	Stability	How well does the application performs with a large amount of data	WP4 & WP5	Internal review, data analysis, log files
	Connectivity	Does the application succeed in connecting to external systems?	WP4	Internal review
Data mining & knowledge discovery	Reality Mining	How well is the application able to capture experiences on urban space	WP1 & WP5	Internal review, personal contact
Usability	Learnability	How easy is it for users to accomplish basic tasks the first time they encounter the design?	WP5	questionnaires, personal contact, app store reviews
	Utility	Does the prototype meet the given requirements	WP5	Evaluation sessions, questionnaires, internal review
	Memorability	When users return to the design after a period of not using it, how easily can they reestablish proficiency?	WP5	questionnaires, personal contact, app store reviews
	Satisfaction	How pleasant is it to use the system in general	WP5	questionnaires, personal contact, app store reviews
	Efficiency	Once users have learned the design, how quickly can they perform tasks?	WP5	questionnaires, personal contact, app store reviews

	Feedback & Errors	How many errors do users make, how severe are these errors, and how easily can they recover from the errors.	WP5 & WP4	questionnaires, log files, data analysis
	Feedback & Errors	Feedback from the system to user	WP5 & WP2	questionnaires, personal contact, data analysis
	Reliability	Does the software runs consistently without crashing	WP5 & WP4	questionnaires, internal review
Issue Reporting	Quality	How do municipalities find the issue reports of high quality?	WP5	personal contanct, internal review
	Dialogue extension	How well used is the extended feedback mechanism option?	WP5 & WP2	Data analysis
Impact	General	User numbers	WP5 & WP4	Data analysis
		Usage numbers	WP5 & WP4	Data analysis
		Media coverage	WP5 & WP6	Internal review

Table 43: Evaluation aspects and methods in the first Issue Reporting & Reality Mining

5.2 User-experiences and customised prototype evaluation

In this section we present the results of the first field trial of Issue Reporting & Reality Mining. First in 5.2.1 the evaluation of the usability of the Issue Reporting client is given. These results are combined with our extensive appstore user-ratings. In 5.2.2 we will focus on the requirements as set out in D5.1, and check how the requirements for the first trial are experienced. Finally, in 5.2.3 we give the specific input for WP1-WP4 for the next phase of Live+Gov. All the evaluations are done from a functional users perspective.

5.2.1 Evaluation of Usability

As stated in the evaluation process (section 5.1.3), to gather the different insights for usability, a smaller number of respondents than might be expected is needed. Although 7 respondents for our questionnaires is rather low to show statistical evidence, combined with our years of user feedback and analysis, it gives the insights needed for our evaluation. For the Issue reporting and reality mining, we mainly focus on quantitative analysis. However, the usability insights on our prototypes are not ignored. For the actual in-depth feedback of users and stakeholders we conducted surveys and interviews to measure the functional usability of the system. These surveys were distributed at the City Safari event (see 5.1.1.2) and were done in order to see if any specific usability aspects for our mobile application should be improved. Apart of this, we did a full analysis of user experiences on appstores, Twitter and emails, for which we have an extensive backlog history for the Issue Reporting client in the production environment: BuitenBeter.

In Figure 17, the overall results on usability aspects of the BuitenBeter Issue Reporting client are presented as measured at the specific City Safari. In next sections, we present more information on each usability aspect.
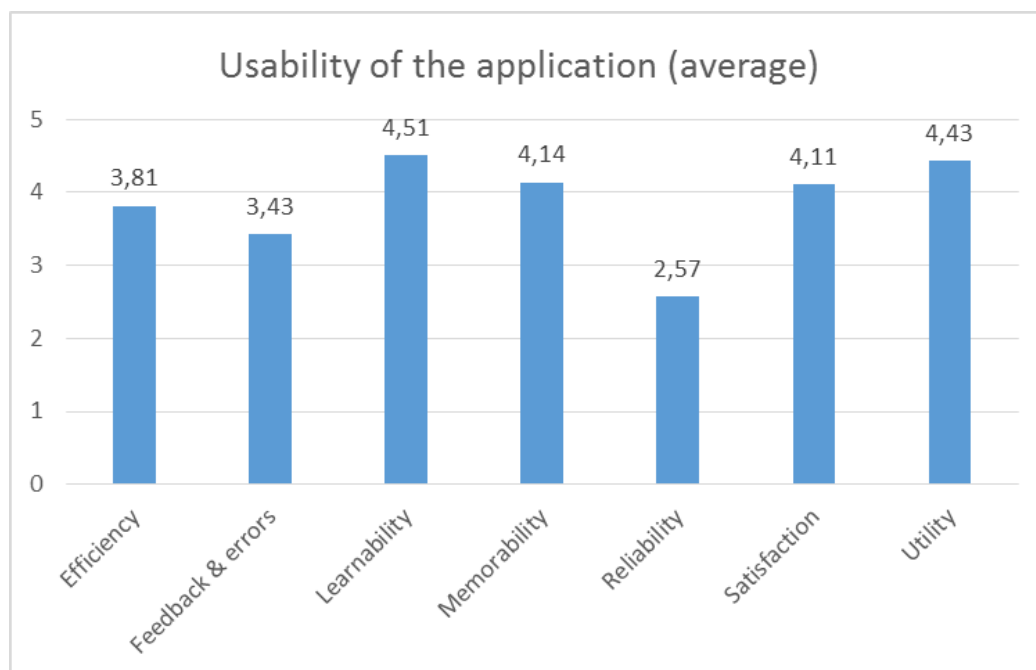
Figure 17: Overall survey-response on usability for Issue Reporting client

As a full usability and user experience check, we have the overall ratings in different appstores (Apple appstore, GooglePlay (android), Windows Phone Marketplace, Blackberry App World) to give a good impression on the status of the Issue reporting client. In total this

data consists of 461 reviews, and the data is normalised for the different appstore rating systems (5 = maximum score, 1 = minimum score).
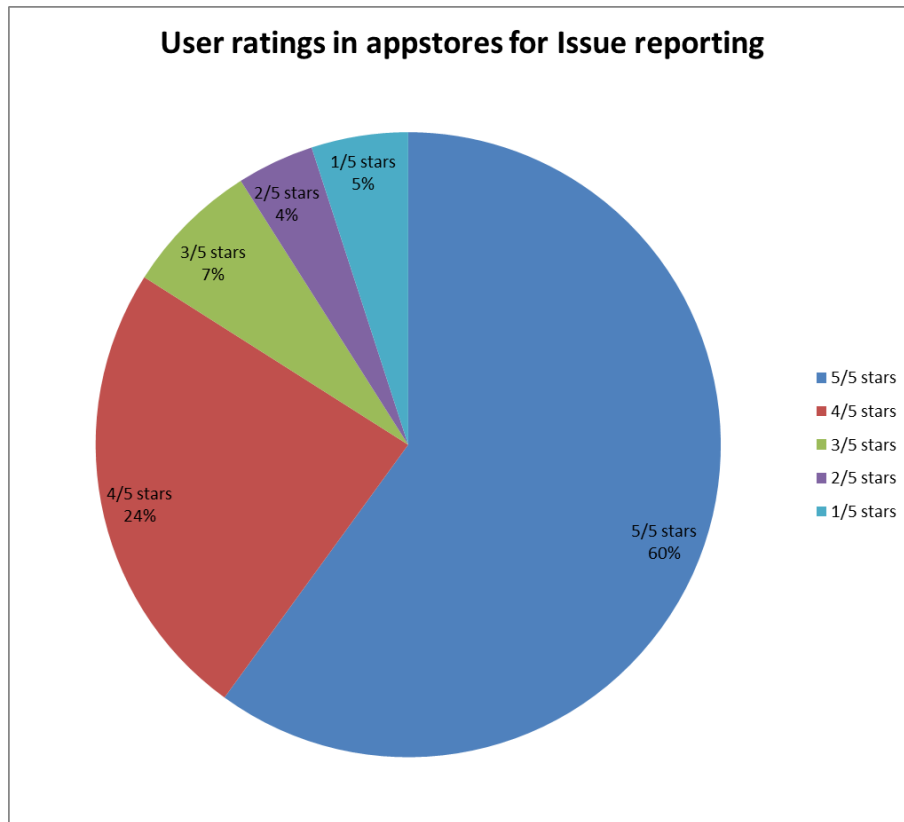
Figure 18: User ratings in appstores for Issue Reporting client (5 = maximum score, 1 = minimum score)

These Figure 17 and Figure 18 show the very positive user-experiences: 60% gives the highest rating possible and 24% rate it with "4/5 stars". All reviews below Neutral only encompass 16% of all the application review-ratings.

5.2.1.1 Learnability

Learnability is measured along different aspects of Learnability. In following Table 44, the results of Learnability and its sub-aspects are given per aspect.

Question topic	Mean	Median	Standard Deviation
Learnability overall	4.51	5	0.51
Easy to learn	4.86	5	0.38
Easy to navigate	4.43	4	0.53
Clear language & wording	4.57	5	0.53
Easy to find information	4.43	4	0.53
Use without written instructions	4.29	4	0.49

Table 44: Descriptive statistics for Learnability for Issue Reporting

These scores leave not much to question on learnability and confirm the overall positive rating. The application has been designed and developed strongly from the citizen perspective, and this shows itself beneficial in the user experience feedback.

5.2.1.2 Utility

Important aspect for the usability is that the system actually follows the expectations of the users regarding what can be done with the application. In our questionnaire surveys, we found the following user-experience on "I can do what I expect with the Issue Reporting client". The results in Table 45 show again a clear positive result.

Question topic	Mean	Median	Standard Deviation
Utility overall	4.43	4	0.53

Table 45: Descriptive statistics for Utility for Issue Reporting

5.2.1.3 Memorability

On the aspect of memorability we measured specifically what the user experience is on the organisation of information. The results of the survey questionnaire are shown in Table 46.

Question topic	Mean	Median	Standard Deviation
Memorability overall	4.14	4	0.77
Clear organization of information	4.43	4	0.53
Results are conform expectations	3.86	4	0.90

Table 46: Descriptive statistics for Memorability for Issue Reporting

Again, this shows a strong positive result and leaves not much room for interpretation.

5.2.1.4 Satisfaction

A sort of overall impression could be seen as the aspect of satisfaction. In Table 47 the results are presented.

Question topic	Mean	Median	Standard Deviation
Satisfaction overall	4.11	4	0.74
User-friendliness	4.14	4	0.77
Design & layout	3.86	4	0.69
Recommend to friend	4.29	4	0.76

Table 47: Descriptive statistics for Satisfaction for Issue Reporting

Here, we see the first non-overly positive results. On the lay-out and design, the focus has always been to have a functional application, for which the lay-out is of secondary

importance. That said, the application is meant to 'shine' on its own as well. These results show that there is room for improvement, like a layout update.

On User friendliness, we see one respondent that experienced this negatively. In the additional comments this respondent made in the survey, it became clear that an important point of critique was that there are no possibilities to "withdraw" the issue report. The respondent regarded this is not-user friendly. This is not a completely new insight for us, but to make adjustments is very tricky. Once the report is sent, the municipality processes start to work. To withdraw the report, the issue-reporting client must take over to the governmental processing of the report itself, which is not the role for the Issue reporting client. However, in the client, we took great care in the right communication to the end-user: before sending the report a specific dialogue pop-up, that must be confirmed by the user, is shown in which it says that the report is going to be sent to municipality xyz. (for more on this, see 5.2.1.6).

5.2.1.5   Efficiency

Efficiency tells us how much time after getting used to the application, is needed to perform tasks. In our surveys we asked for the user-experience on three sub-aspects, with the goal to give us insights on how the efficiency is experienced by users. In Table 48 the results are presented.

Question topic	Mean	Median	Standard Deviation
Efficiency overall	3.81	4	0.98
Report made in few steps	4.29	4	0.49
Consistency	3.29	4	1.25
Straightforwardness	3.86	4	0.90

Table 48: Descriptive statistics for Efficiency for Issue Reporting

For efficiency, we had clear miss-match with the question asked, and the information we wanted to get. On the topic of consistency, we can actually see in the free-comments on the surveys that all respondents experience the application as coherent and 'logical'. However, we do see negative feedback on the consistency. After further inspection of this clear mismatch, we found the confusion. The question was interpreted not in the sense whether or not the application is coherent and logical, but weighted the fact that the devices were getting resource low. In the City Safari, the use of the application is slightly different than usual: instead of making the report, it was the case to make a lot of reports. This means that the application never closed and that non-stop GPS service was active. The results shown for consistency are skewed by our mistake to present an unclear question. This is an issue for reliability and not for efficiency.

5.2.1.6   Feedback & Errors

For usability, we also evaluate the aspect of Feedback & Errors. By this, we mean the feedback the system provides to the user in terms of actions that are available or incorrect and the options it serves for correcting user-errors. Now, all these usability aspects are of course not fully exclusive. When on one aspect a user has a negative experience, the chance that this influences another aspect is rather high. Here, we see overlapping results for the Feedback & errors and the User friendliness as described for Satisfaction in 5.2.1.4.

Question topic	Mean	Median	Standard Deviation
Feedback & errors overall	3.43	3	0.94
Possibilities are clear	3.43	3	0.98
Easy to recover mistakes	3.43	3	0.98

Table 49: Descriptive statistics for Feedback & Errors for Issue Reporting

Here, we talked about the (non-existent) possibility of withdrawing a report. The same issue has been reported in the survey for Errors & feedback, which is understandable and justifiable. The solution we implemented is to have a clear dialogue pop-up before sending reports.

Interestingly, when we evaluate with our connected municipalities, the quality of the reports is often explicitly praised. What does this mean? Users might feel uncomfortable not having the possibility to withdraw a sent report. However, this negative experience seems to be rather theoretical as in practise, we have had almost no feedback of municipalities of bad-quality reports. Also, when evaluating the reports ourselves, we see there is only a handful of obvious "joke reports" and again the system is actually praised for its high-quality reports. The results we see here, we do not weigh too heavily as they are more a theoretical experience, we have a lot of supporting information that there are hardly any "unwanted reports sent". On top of this, in our opinion, we implemented what is possible to take great care of high quality reports with the specific dialogue as shown in Figure 19.

Figure 19: Confirmation dialogue in order to maintain high-quality Issue Reports and prevent mistakenly sending

5.2.1.7 Reliability

The stability or reliability of the system as experienced by the users is presented in Table 50. The event had a specific goal, which was to make a lot of reports and have the application active non-stop. This is somewhat comparable to our data-collection for *Jij Maakt Utrecht*, see section 4. However, for the data collection we made specific adjustments to the application, which we didn't make for the City Safari. This showed in the reliability. The resources usage is handled by the operating system, but this of course effects the application as well. What was experienced was that the issue-reporting client often asked the user to turn the location providers on for setting the location. Leaving the resources on nonstop (which happens with non-stop reporting without closing the application), made some phone models to completely run out of battery and turn themselves off.

Again, this is a reliability issue that isn't reflecting the issue-reporting client by itself. It just shows that for specific usage as in the City Safari and the mobile data-collection for *Jij Maakt Utrecht*, adjustments should be made, especially for the GPS-sensor. These adjustments had been made for the intensive data collection, but weren't made for the city safari; which explains the reliability score. For transparency reasons, we do share these *wrongly* measured results in Table 50.

Question topic	Mean	Median	Standard Deviation
Reliability overall	2.47	2	1.13

Table 50: *Wrongly measured* descriptive statistics for Reliability for Issue Reporting

5.2.2 Evaluation of the requirements

After the evaluation on usability aspects, we evaluate each of the functional requirements for the first cycle of the prototype. For each implemented requirement for the first trial, we present the evaluation results and insights.

5.2.2.1 Evaluation per requirement

L+G-FR.02	Be able to sense the context of users (e.g. location, means of transportation, activities etc)

With the issue-reporting client, we put great emphasis on the reality mining aspect: citizens-as-sensors approach. Completely designed from the citizens' perspective, the client makes it possible to send reports on urban space related experiences. In our City Safari scenario, we took a different approach following the same philosophy. For this 1 day event, the users were able to capture reality and provide it with the exact location, their experience of it with customised tags, photograph, and textual description. The Issue Reporting client had the customised tags dynamically available for the specific Jekerkwartier area, in the same production BuitenBeter client as is available for the rest of the Netherlands.

The results of the context capturing for the issue reports is regarded as very effective. Municipalities praise the quality of the reports and the insights they give on the experienced urban space. The results of the City Safari were also met very positively by the involved stakeholders and led to inspiring discussions on the subjectivity of experiences of our live environments.

In current state, we haven't assessed importance to the HAR functionality of WP1.This comes from both a functional level and a risk-management approach. As we have one production client for the entire Netherlands, we aren't in the position to experiment. Functionally, we haven't been able to assess the added value of inclusion of means of transportation or activity. We will have to do the functional assessment now the results of the Mobility trial are evaluated.

Overall requirement evaluation and refinement

L+G-FR.02 has been met successfully; the quality of the Issue reports is regarded as very high. For short-term cycles, the information provided give municipality great insights on specific Urban Maintenance issues. On the more abstract level of citizen experiences, the mobile client provides municipalities with a large amount of information of the experienced state of the real life-worlds. The requirement links to L+G-FR.10 and L+G-FR.19 as well.

Table 51: Evaluation of L+G-FR.02 in Issue Reporting

L+G-FR.05	Provide communicative tools for government and citizens (including reactions, alerts, messages)

The requirement has been addressed with the advanced feedback mechanism. This mechanism is dynamically available for connected municipalities (14 in our trial). The mobile client is the same for the entire Netherlands, the option to provide textual feedback is centrally organised. When the municipality provided textual feedback, this was visible for the citizen who made the report.

The requirement was technically not the most complex feature. However, the implementation on a running system as widely used as BuitenBeter was a challenge. Furthermore, the option has been built in all our supported platforms:

- iPhone
- Android
- BlackBerry
- Symbian
- Windows Phone
- Windows Mobile

The implementation went very smooth and no problems for the application have been stated.

With the advanced feedback mechanism, the communication loop has been truly extended as it provides municipalities with the possibility to communicate their policies and the planned solutions for the reported issue. This is of course a rather fundamental extension for the citizen as well: it provides them with insights on the governmental side of their experiences.

The numbers of provided feedback have been real successful: for more than 70% of the reports textual feedback has been given. In absolute numbers this means that in our trial, citizens got textual feedback for 14.291 reports in total.

We've checked whether or not the extended feedback made the frequency of reports rise per user. For this, we compared our trial data with data of our database for our non-trial population. At current moment, we cannot report on a statistical correlation. For the next phase of the project, we will look into possibilities of fine-tuning the mechanism in terms of visibility of new updates (possibility is push-mechanism notifiers). Also, we are investigating possibilities of feedback in form of a photograph (for example a photograph of the situation after the municipality worked on it).

Overall requirement evaluation and refinement

L+G-FR.05 has been met with great success. 14.291 reports were accompanied of textual feedback by our 14 trial municipalities. The implementation on our running production system went smooth, even with the support of 6 different platforms (iPhone, Android, BlackBerry, Symbian, Windows Phone, and Windows Mobile).

Future exploration will be done in possibilities of photograph feedback and notifier options.

Table 52: Evaluation of L+G-FR.05 in Issue Reporting

L+G-FR.08	Provide possibilities for textual feedback. For example how his contribution is taken into consideration, including possible benefits for participating and status report on the situation
This requirement is directly linked to the extension of communication possibilities as described in L+G-FR.05. The implementation of the textual feedback mechanism has been met successfully and the feature has been used extensively in the trial.	
Overall requirement evaluation and refinement	
More than 70% of the reports in our trial have been provided with textual feedback (14,291 reports in total). L+G-FR.08 is met very successfully. The option has been praised by both citizens and public officials. For next steps we will consider notifiers (push mechanism could be a possibility) and the possibility for municipalities to provide the reporter with a photograph (of the situation after it has been dealt with).	

Table 53: Evaluation of L+G-FR.08 in Issue Reporting

L+G-FR.09	Allow a user to view all of the issues that are currently open, and also a history record of "my issues" (where the citizen has participated) and "my municipality" (where all of the past issues from the town can be found) with a short summary of each of the past issues.
This requirement reflects one of the core functionalities of Issue Reporting. Apart from making reports, it is needed to be able to see the filed reports. In the issue reporting client, the user is able to see a list of his made reports. For each report, detailed information is available. The exact location (on a map), the photograph that was taken, the status of the report and when available: the textual feedback of the municipality. In its current state, there are no possibilities for ordering the reports. We are careful that the application stays user-friendly and aren't overloading users with options, features and possibilities that distract from the purpose of the application and confuse the user.	
Overall requirement evaluation and refinement	
L+G-FR.09 is met successfully: the user is provided with a clear overview of his reports, with possibilities to look for further details for each of them.	

Table 54: Evaluation of L+G-FR.09 in Issue Reporting

L+G-FR.10	Be able to utilize 'citizens as sensors' for their environment by mining relevant data (e.g. location, context, pictures, reports) from their reality
With the Issue Reporting client, citizens are provided with possibilities to report urban space experiences in an easy manner. This requirement is the core idea for Issue Reporting in Live+Gov. The mined reality by using citizen as sensor approach exploits a crowd sourcing technique to generate better understanding of public space. This way, municipalities are provided with insights for the Urban Maintenance domain at low cost and in an encompassing manner.	
Overall requirement evaluation and refinement	
L+G-FR.10 is one of the crucial requirements of the Urban Maintenance Use Case. Within the trial close to 20.000 detailed reports have been made, of which 15.357 came with a photograph (77%). All reports have the exact location and around 85% are accompanied by a user description. These reports, as a result of citizens as sensor techniques, give tremendous information to municipalities. L+G-FR.10 is successfully met.	

Table 55: Evaluation of L+G-FR.10 in Issue Reporting

L+G-FR.13	Be able to connect to different municipality administration systems
For our trial municipalities, we can report that we have a connector with the second largest municipality (Eindhoven with 218,433 inhabitants). We have a standardized StUF connector with the administrative system of the municipality. Handled by the connector are status updates and the distribution of the textual feedback to the citizen. StUF is the standardized connector for the Netherlands. The connection is fully implemented and with our trial we prove our architecture being able to connect on public administrative systems. See also D4.2 [6].	
Overall requirement evaluation and refinement	
Within our trial we have proven with the StUF connection of the Issue Reporting client with the backend system of Eindhoven that we're able to interact with external administrative systems. L+G-FR.13 has been successfully met in our trial. This success is especially relevant for WP4 in which the possibilities to connect with public systems is located.	

Table 56: Evaluation of L+G-FR.13 in Issue Reporting

L+G-FR.19	Collect and mine citizen experience and citizen (implicit) norms
Closely related to L+G-FR.10, within our Use Case we aim to mine reality and citizens experiences. Our Issue Reporting client implicitly collects citizens' norms about urban maintenance. Moreover, with our Jekerkwartier City Safari event, we specifically mined negative and positive experiences about the urban space. With customized tags (locally dynamic), we used the Issue Reporting client to gather the subjective feelings about citizen-chosen locations. With the possibility to also rate positively, a nice variety of different opinions and implicit norms about locations were gathered.	

Overall requirement evaluation and refinement
The conceptual change as experimented with for the City Safari has been interesting. Whereas with our normal issue reporting, the collection of citizen experiences is focussed on 'issues', in the City Safari people were able to send 'reports of praise' as well. For future steps, we see opportunities to explore the possibilities of positive ratings. Linked to the mining experiences and implicit norms, we want to explore the possibility to communicate existing service level norms. This would inform the citizen greatly, and would make it possible to match the experienced norm and the official quality level. With these possible future steps in mind, we conclude that L+G-FR.19 is met successfully.

Table 57: Evaluation of L+G-FR.19 in Issue Reporting

5.2.3 Evaluation results for specific WPs

5.2.3.1 WP1 – Reality Sensing and Mining

At current state, no components of WP1 are used in the Issue Reporting field trial. However, the Use Case is focussed on reality mining, which is an important aspect for WP1. With our mobile client, we collect citizen experiences of urban space. Interesting future developments rely on the possibilities of the reality mining and pattern recognition possibilities provided by WP1. These knowledge discovery technologies are planned for upcoming period in WP1. From our field trial we want to re-state the importance of the pattern recognition aspect.

5.2.3.2 WP2 – Policy Modelling and Visualization for Mobile-aware Dialogue

With the extended feedback mechanism, we extended the mobile dialogue between citizens and municipality. We provided the municipality with the possibility to communicate their policies and processes to the reporter. By the citizen-as-sensor techniques, we offer the municipalities with detailed user experiences. Those two facets combined are assessed as crucial for a next level citizen-government dialogue. Future development will take the communication of service level norms in consideration as well. The successes booked of extending the dialogue and creating a feedback loop between citizen and government are important for the three-pillar strategy of WP2 in the transition towards we-government policies. With the field trial, we proved that we made considerable steps towards a more transparency and more citizen participation. In future WP2 work, special attention should be given within the Training Package for combining the lessons learned in our Use Case for future participation initiatives.

5.2.3.3 WP3 – Mobile Augmented Reality

Related to our extension of the feedback loops of citizens and governmental bodies is the functionality located in WP3. Personalised content delivery in combination with data mining techniques of WP1 would provide possibilities of offering state of the art context aware feedback tailored to specific citizens' needs. In the upcoming period, we will need to strengthen the linkages on this aspect with WP3.

More important is our requirement to provide citizens with the service level norms for urban maintenance objects. The envisioned mobile presentation methodology is augmented labels in the Issue Reporting client. The augmented reality functionality has been tested within the Urban Planning Use Case and we'll use upcoming period to assess in close relation with WP3 how to implement the foreseen functionality in the Urban Maintenance Use Case. At current

state, we restate the importance of the offered functionality by WP3 for the Issue reporting and reality mining scenario.

5.2.3.4   WP4 – System Integration

Public Administration Connectors

An important aspect of the Issue Reporting Trial is the connection to the Public Administration Systems. During the trial two kinds of connectors are used to exchange issue reports and feedback between the issue reporting tool and the public administration. The cities of Eindhoven and Rotterdam use an advanced StUF connector [6], which follows the Dutch national standard for exchanging electronic forms. The other municipalities receive an e-mail with the issue report information and give feedback via an online web form. Citizens receive the feedback on their issue reports in the mobile application itself. In some cases an additional e-mail is sent by the Public Administration System.

This automated exchange of information between the Live+Gov solution and the Public Administration Systems proved to be very effective and reliable. The StUF standard is matured and implemented by the most major suppliers of Public Administration Systems. Technically no major problems are encountered during the trial. A few issue reports could not be forwarded automatically because they exceeded the maximum photo size accepted by the Public Administration System. Due to the automatic monitoring procedures these errors were detected automatically. The photographs were resized and forwarded manually. For the 2nd trials an automatic resizing mechanism will be added.

Operationally more alignment is required between the Issue Reporting Tool and the Public Administration systems. For example, in some cases a confirmation mail is sent by both the Issue Reporting Tool and the Public Administration System. Some users experience this as spam.

Furthermore ideas for more advanced feedback options are raised. In the current form the feedback is limited to the issue report status and a free feedback field. It would be interesting to add more information like targeted resolution times according to the local policy or photographs of the situation after the issue is solved.

Scalability / Stability / Availability

The issue reporting trial is performed in the BuitenBeter production environment, which gives access to the existing BuitenBeter user base, but also raised the requirements for stability and availability. Thanks to the extensive test procedure and o performance issues or downtime was experienced during the trial. Section 5.2.4 gives more insight on the number of issue report and the amount of feedback given during the trial.

Cross-platform

During the trial issue reports were filed with the issue-reporting tool on multiple mobile operating systems. BuitenBeter has native clients available for iPhone, Android, Windows Phone/Mobile, Blackberry, and Symbian. These Issue Reporting clients are used and extended with the option to present feedback to the citizens. This feedback option is activated or deactivated depending on the location where the issue report is filed.

5.2.4   Impact and success of the Use Case trial

With the Issue reporting Use Case, we have the following figures for our trial. In Table 58 total amount of reports and unique trial users are presented. We also present the numbers in Figure 20. These numbers encompass the reports made in our trial population: the 14 municipalities as described in Table 38 in the trial set-up section 5.1.1.

Topic	Number
Total issue reports made during the trial	**19,877**
Total issue reports *with* textual feedback from municipality	14,291
Total issue reports *without* textual feedback from municipality	5,586
Unique users that made issue reports in the trial	**5,279**

Table 58: Issue Reporting trial: impact in numbers

In Figure 20, we present the key numbers regarding the impact we made in diagram:
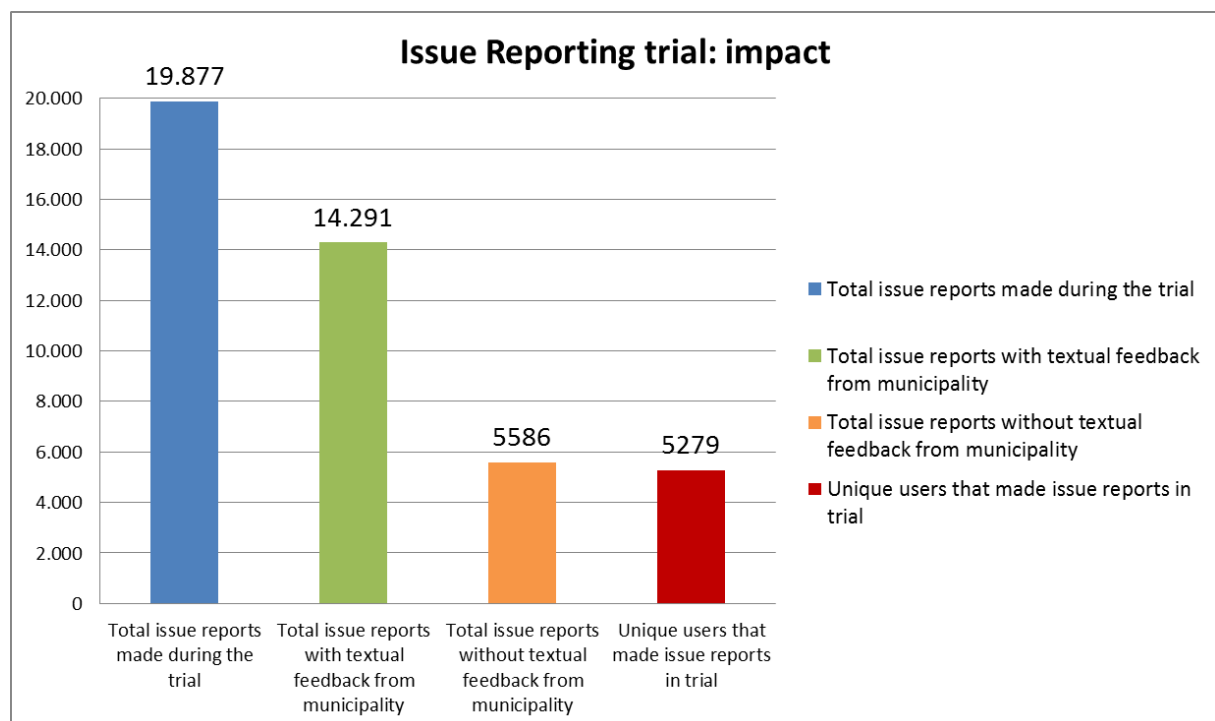
Figure 20: Issue Reporting trial: impact in numbers

From these figures, we see that more than 70% of the reports made were provided with textual feedback from the municipality. The distribution of reports with and without textual feedback for our trial is presented in Figure 21.
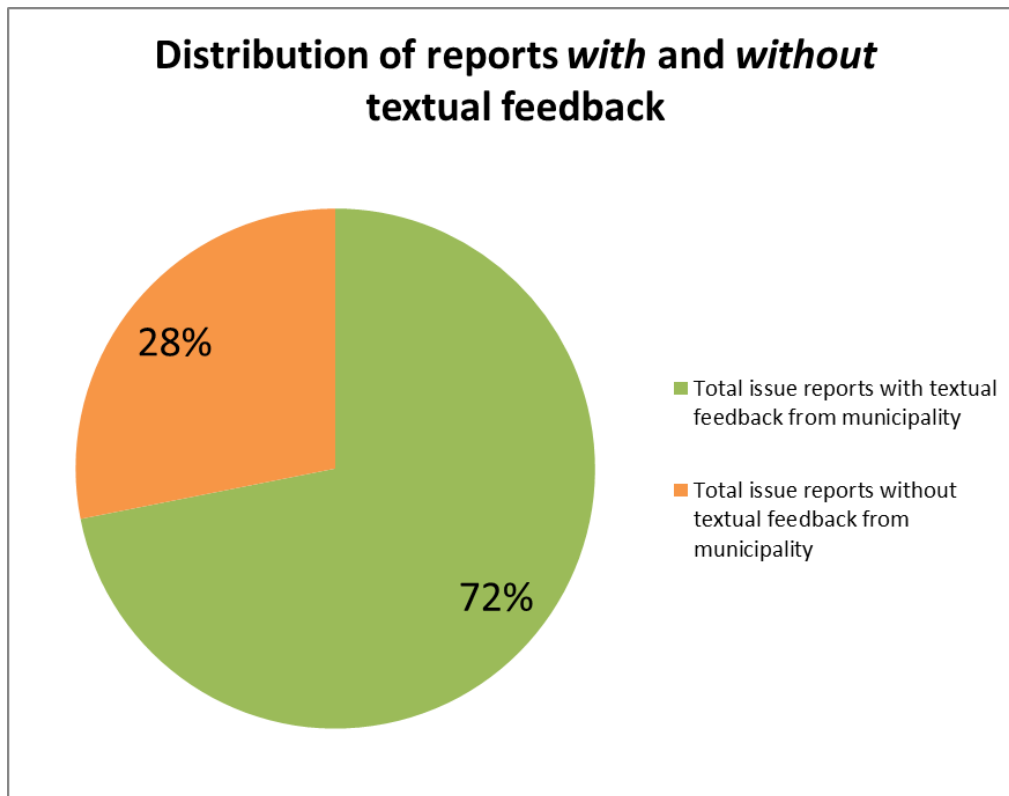
Figure 21: Issue reports with and without textual feedback

5.2.4.1   Jekerkwartier City Safari

For the purpose of sharing citizens' experience, four extra categories were added to BuitenBeter specifically for this area. Two for positive value: Super Top and Top, and two for negative value: Flop and Super Flop. As can be seen from Figure 22, about half of the 127 reports reflect a positive experience and the other half was negative. The participants were as happy to share their positive experiences as much as their negative experiences.
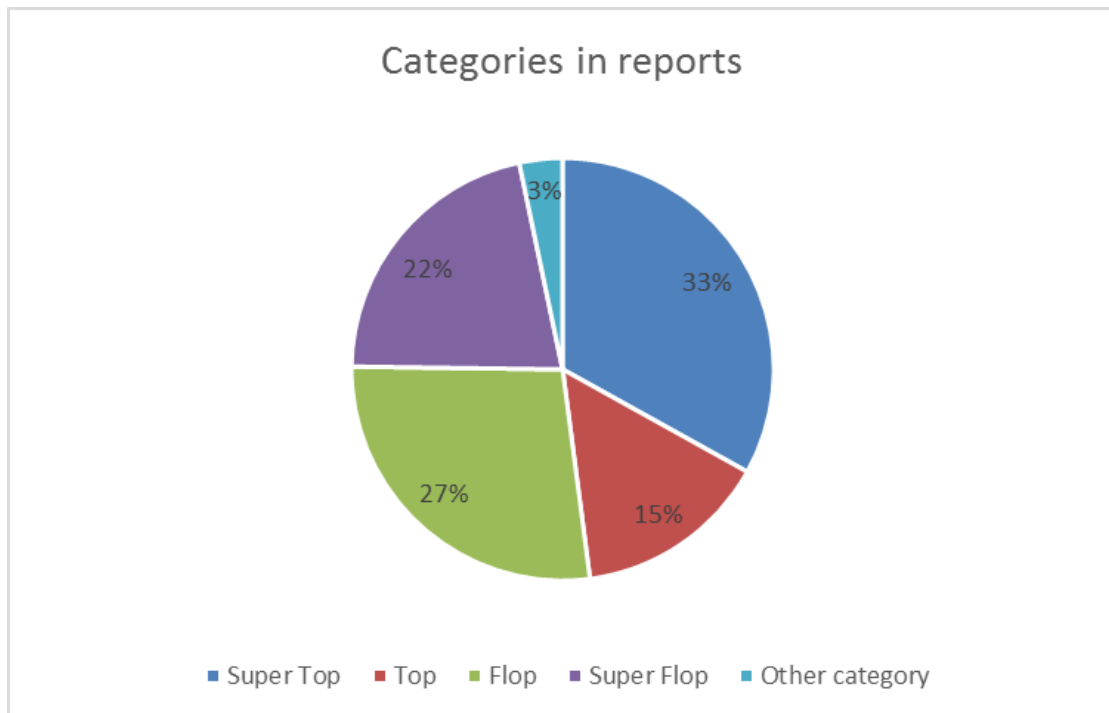
Figure 22: Categories of the reports that were made in Jekerkwartier, Maastricht

Eventually, a total of 13 different smartphones were used that covered iPhone, Android, BlackBerry and Windows Phone devices. On average, each trial user filed 9 experiences during the City Safari.

Publically, the trial appeared on a regional news website, as can be seen in Figure 23:

Figure 23: Article about the City Safari trial on regional new website

Furthermore, important impact has been made on a political level as well. In a recent public debate (February 8th, 2014) for Maastricht for upcoming local elections, the Issue reporting client, the web visualisations of the reports and the City Safari itself got full attention and praise by all participants of the debate.

5.3   Summary of the evaluation of Issue Reporting

With the Issue reporting and reality mining trial, a large-scale implementation of the customised Issue Reporting application has been realised. More than 19,000 reports have been sent by over 5,000 users to our 14 trial municipalities. For over 70% of these reports the citizen has been provided with textual feedback. The evaluation of the experienced usability is positively met in our trial and in the application reviews of appstores (over 80% positive to very positive).

Next to this large-scale implementation, also a smaller-scale one-day event has been conducted: a City Safari in Maastricht. In this closed user group setting, reality mining with the usage of the Issue Reporting client has been tested. This has been done by a customised client that provides the user in this area with dynamic tags. The experience mining during this event has been rated very positively by both of the end users as well as on the organisational level. The gathered experiences and the aggregated visualisations of the data immediately opened up dialogue for the stakeholders. This catalysing element of the visualisations has been praised by involved participants. Important to note is the recruitment of participants for trials. In situations like the City Safari, which involve a specific event, it turns out that recruiting users shows itself to be challenging. Initially, the organisation of the event targeted for 50 participants. In the end the event involved 15 trial users. This is seen as a positive result. The lessons learned from this provide valuable input for the Training Package that is being developed, as it provides insights of implementation. Activating new initiatives asks for a large amount of effort to attract participants on behalf of the local government.

The requirements have been met successfully. This is expected, as this trial has been conducted on a production system that is running for the entire Netherlands for over 4 years. With this amount of experience, realistic prognoses can be made on expert knowledge of the Urban Maintenance domain. The quality of the reports, and mining citizen experiences and implicit norms is regarded of high quality by municipalities. This is due to the citizen-as-sensors philosophy: by leverage of citizens to mine reality, real detailed information about urban space and how it is experienced can be collected. The Issue reporting client and the extended feedback mechanism are deployed on 6 different platforms (iPhone, Android, BlackBerry, Symbian, Windows Phone, and Windows Mobile). Over all these platforms, the architecture proves itself stable for large amount of usage. The option is dynamically available based on location; a strategy that fits our SaaS based approach neatly. With the standardized external connector (StUF), we are able to connect directly to administrative systems whenever municipalities are equipped for the automatic connections.

For future development, possibilities to make the provided feedback more visible will be investigated. An example of this is current research about notifier options in form of push-functionality. Another possible future development can be the extension of the Issue Reporting client with an augmented presentation of service level norms. These extensions will contribute to strengthened transparency and citizen-municipality dialogue. Last but not least, with work of WP1 pattern recognitions (data mining); focus will be on presenting advanced visualisation of the reported data. This will be valuable input for municipalities to adjust their Urban Maintenance strategies to citizens' needs. A close relationship will be

maintained between Use Case owner Yucat and the related WP1, WP2 and WP3 to bring these aspects of Live+Gov together.

6 Urban Planning Use Case: Urban Planning Gordexola

Before going into detail about the first field trial of Urban Planning Gordexola we need to stress the importance that this Use Case gives to the engagement of the decision makers. Their commitment to this conceptual change is absolutely necessary in order to succeed in the aim of opening modern communication channels between citizens and administration in the political arena of public planning. Giving people power in public plans has a great potential: it is becoming a social demand that at the same time is contemplated in legislation, as has been deeply depicted in D2.1 [5]. Sharing their power to decide with citizens on one hand supports the decisions made by administrators, but it is also tricky and somehow risky, making the people in charge of the municipality hesitant towards this type of initiatives. This Use Case tries to use ICT to help accelerate the changes in order to allow eParticipation and enhance the public planning at the local level. At the same time, if the participation process is not used for final decision-making or citizens do not receive feedback it can be counter-productive.

Therefore the main goal of the first trial is acquiring the trust of politicians and their acceptance of the system, as well as their commitment to fully implement the eParticipation process proposed during this First Trial in a bigger scale in the Second Trial. First a common understanding about the aims and objectives of the participation process need to be clear: the potential political outcome, its potential political leverage and impact. Opening a fake participation process can have the effect of distrust, scepticism and cynicism towards administrators, as stated above, and also detailed in D2.1.[5]

The following sections in this section will describe the first trial of Urban Planning Gordexola and the prototype application that carries the same name. This mobile application for citizens which was already detailed in D5.2 is now used on field for the trial. [1] Finally the results of the evaluation of the prototype and user experience and the impact of the trial will be further looked into.

6.1 Urban Planning Field Trial

6.1.1 Field Trial set-up

The Field Trial of Urban Planning is the trial of a participation process in Gordexola by means of mobile devices, that in general can be classified according to the following differential criteria, that was included in D2.1 [5].

As for the organizational features, the participation in the plans in this Use Case have a punctual duration as they are related to a plan and therefore have a closing date, where the numbers collected are considered final. The size of participants will vary as the participation planned for the next trial will be open to a wider scope of citizens. In the first field trial three plans were presented to a group of 12 citizen users. The duration of the participation was short due to the procedure that was followed.

The recruitment of participants will be self-selection in the overall field trial, complemented with group dynamics with associations and active groups in the municipality, as the aim of the Use Case is to have a broad reach. This is a cornerstone of the project as the intention of using the mobile device to enhance citizen-government communication and allow citizen participation is exactly this, to reach more people, initially not interested in the traditional

ways of taking part in the political decision making process. However for the first field trial a targeted selection has been made in order to have first reactions from a group of trust on the developed system and avoid technical mistakes having a negative impact on the long-term trial and as a result of the Use Case.

The forms of communication included articulating interest, negotiating and exchanging arguments. The application itself articulates interest by means of sharing information of interest related to the plan, the voting represents the negotiation, and the exchange of arguments is through the "other comments" option, the real time results screen, and in the future version in the "final decision feedback".

For the first trial, the communication through the application was complemented by a group meeting with the selected users, who voluntarily accepted to join, on the 20th of January, 2014. In this meeting trial users received more information about the trial, the plans, the information presented, putting into practice the three forms of communication that the Use Case contemplates. For the second field trial these meetings will be held with the target-selected groups to complement the communication that goes on directly through the mobile application.

The Urban Planning Gordexola application was developed by YUCAT, with a great involvement of CERTH, and the user recruitment and support was the responsibility of BIZ. The application was ready and handed over to BiscayTIK after the integration work was finished at the end of November and then the on field system testing was conducted.

As for the aims and objectives, the urban planning field trial is clearly intended for co-deciding and co-governing. The combination of ICT, mobile technology in this case, is intended to extend the scope of the audience that can be reached and therefore involve more people during a longer period of time and increase the representativeness of the opinion collected in the process. The potential has been demonstrated through the first Trial to the Town Council of Gordexola, to the elected representatives. In this first phase the results have been presented as reports for the decision makers to use in their current decision making process, without more detail about how the consultation results may be used by them. Work in this line will continue in order to have a clear idea about how the results will be integrated in the formal political process. Moreover, in order to avoid distrust, as mentioned before, and have more commitment of citizens who are interested in public plans, for the second field trial this will be further studied and details will be presented to the citizen participants, making the aims and objectives of the participatory process clear from the beginning. Citizens will be informed of the consequences and purpose of their participation in order to avoid exaggerated expectations.

Therefore the First Urban Planning Field Trial consisted in the presentation of the work developed in the Use Case and detailed in D5.1 [2] and the prototype from D5.2 [1] to the stakeholders. The main stakeholders for the Urban Planning Use Case are the citizens of Gordexola, and the Town Council (in this case the decision makers are the Councilors who are members of the governing body).

As a first step the meeting with the group of 12 selected citizen trial users was held on the 20th of January, 2014. The gathering took place in the Town Hall of Gordexola as the weather made it impossible to be outdoors on all the sites all the time as it was planned initially. This was possible thanks to the participation of Iñigo Gómez, a member of the Advisory Board, and active collaborator in the project, who works as a civil servant in this town. He was also

in charge of selecting the trial users from Gordexola, doing a targeted selection to reach people concerned with public issues and representative of the community in Gordexola who were willing to participate in this first trial phase. The group included people of diverse type, with ages going from 15 to 55, including males and females, and both people interested in technology and people who are not.

Figure 24: Urban Planning Gordexola- Citizen meeting

During this meeting the Live+Gov project in general and the Urban Planning Gordexola Use Case in particular were briefly presented. Then the application developed for Urban Planning Gordexola, which was part of D5.2, was made available for the trial users, who were able to install it on their smart phones. The application was shared with them via email or Bluetooth, as preferred by each person. They were also handed with the user manual and received an explanation about how the application worked, as well as assistance if they had problems, by people from the BiscayTIK Foundation that were there.

Figure 25: User manual for the 1st trial of the Urban Planning Gordexola application

Then they were offered the chance to use the application on their own, explore the options and see how it worked. The users had the chance to see the Augmented Reality both directly on the sites triggered by location, as well as triggered by image recognition, on the pictures on information posters and on the screen of a computer to resemble how it would look on the municipal web page.

Figure 26: Poster used in the 1ˢᵗ trial of Urban Planning Gordexola for image recognition

After some time given for using the application on their own, and as a second part of the gathering, they were invited to fill in a questionnaire and they were personally asked about their general impression.

Figure 27: Citizen user-experience questionnaire 1st trial of Urban Planning Gordexola

The next step was the meeting with the councillors on the 22nd of January, 2014. The two councillors present at the meeting were Ignacio Gutiérrez del Olmo, who was also the Deputy Mayor at that moment due to a leave of absence of the Mayor Iñaki Aretxederra Zurimendi, who would have also participated, and Juan Ramón Melchisidor. They represent 50 percent of the governing body of the town council of Gordexola, and are the councillors in charge of public infrastructures and new technologies, respectively.

During this get together, firstly the citizen part was presented to the administrators; they were shown how the Urban Planning Gordexola application worked, the different views and possibilities. Then they were presented with the handmade reports containing the results collected through the trial. The reports were very visual to offer the administrators the general information collected in a way they could quickly get insights from. The results were segmented and analyzed based on the information collected from users about age, gender, area of residence, etc. These handmade reports were a representation of what would be offered to them in the second trial as part of the administrators' access to results, together with the information that the council already has in their system to compare with total numbers of residents, age pyramids, etc.
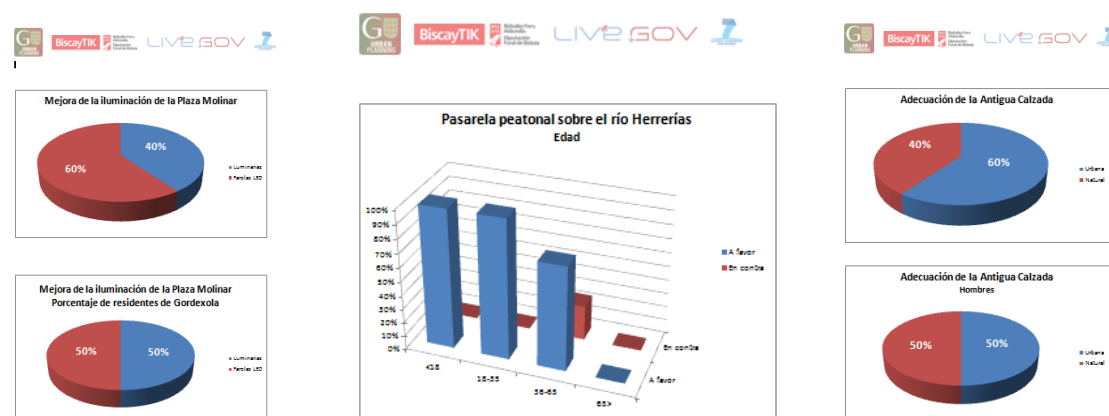
Figure 28: Handmade reports to present how the results could be studied by decision makers
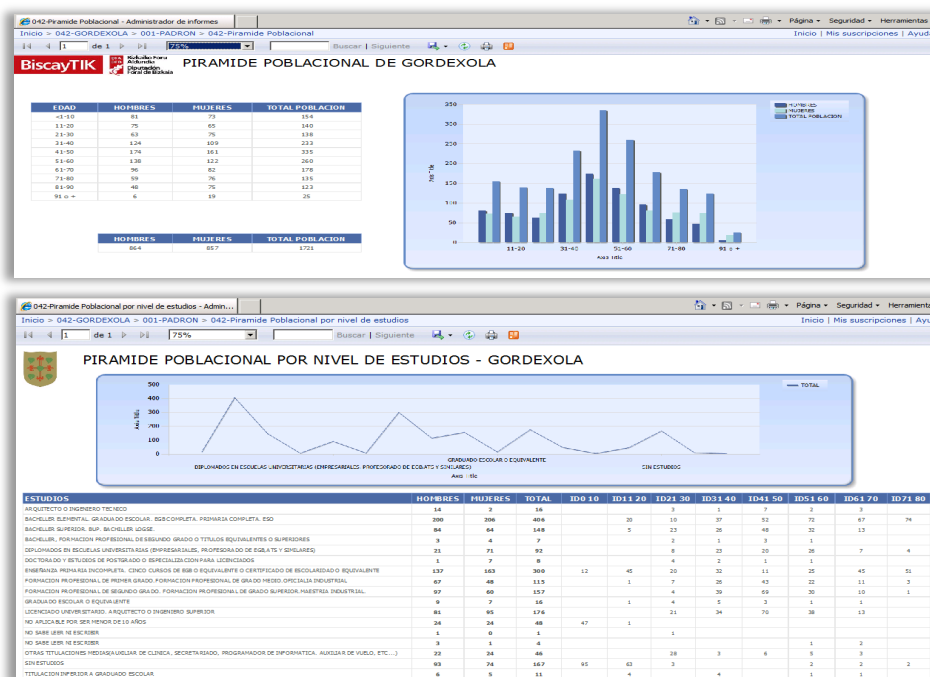
Figure 29: Citizen information presented in the municipal system of Gordexola

This meeting can be considered the first meeting of the so called "Working Group", as in the initial phases of the set up for the first trial, the technical and political plans presented were selected and worked on by an initial working group. This working group is where the relevant issues of the plans, technically and relating to information exchange, will be worked on, as well as the follow up of the process along time. The reason for the initial working group to be different was in order to show the political side a tangible result that would be easier to understand, feel comfortable with, and find of interest for their decision making tasks before their hands on inclusion.

After the first part of the meeting, which consisted in sharing information with them, the second part of the meeting consisted in collecting their impressions through a questionnaire as well as through direct interview questions.

Figure 30: Administrator experience questionnaire 1st trial of Urban Planning Gordexola

Therefore questionnaires and personal impression collection was performed with both main stakeholder groups, citizens and administrators, in order to get insights on their opinion on the quality of the initiative and the usability aspects of the application.

6.1.2 Urban Planning Gordexola mobile prototype

The Urban Planning Gordexola prototype had to comply with the following functional requirements that were included in D5.1 and D4.1.

No.	Live+Gov system requirements
L+G-FR.02	be able to sense the context of users (e.g. location, means of transportation, activities etc)
L+G-FR.03	Be able to show information in a context aware manner (location, person, time-specific, situational) to create greater awareness and create mutual citizen-government understanding
L+G-FR.04	Provide possibilities to group, aggregate, filter large amounts of data to discover knowledge
L+G-FR.05	Provide communicative tools for government and citizens (including reactions, alerts, messages)
L+G-FR.06	Provide possibilities to visualize knowledge and patterns from data by data-mining techniques
L+G-FR.07	Be able to capture images and recognize their content
L+G-FR.08	Provide possibilities for textual feedback. For example how his contribution is taken into consideration, including possible benefits for participating and status report on the situation
L+G-FR.09	Allow a user to view all of the issues that are currently open, and also a history record of "my issues" (where the citizen has participated) and "my municipality" (where all of the past issues from the town can be found) with a short summary of each of the past issues.
L+G-FR.10	Be able to utilize 'citizens as sensors' for their environment by mining relevant data (e.g. location, context, pictures, reports) from their reality
L+G-FR.11	Provide (mobile) presentation possibilities in an Augmented manner
L+G-FR.12	Provide possibilities to retrieve personal opinions, votes or comments from users
L+G-FR.13	Be able to connect to different municipality administration systems
L+G-FR.21	Provide possibilities to communicate urban planning plans

Table 59: Requirements for first Urban Planning Gordexola Mobile trial

The components that have been used for the mobile application for the first trial of this use case are included in the next table that was further detailed in D4.1 [11] and D5.2 [1]. All of these components have been integrated and used for the Urban Planning Gordexola application.

#	Live+Gov component	Notes
C04	eGovernment Dialogue and Visualisation Service	This back-end service is responsible for the storage of the initiatives, questionnaires and the results.
C05	Augmented Reality Presentation Service	This Backend Service hosts and serves the AR information (3D models of the plans, GPS coordinate, meta-tags, etc.)
C06	Object Recognition Service	Backend Service that recognises visual content. In the Use Case it will allow the application to show 3D models when a target object is detected.
C10	SaaS Service Center	This supporting service offers SaaS specific functionality like account management, access control and diagnostics.
C12	Mobile Augmented Reality Client	This mobile client used for Urban Planning is based on Metaio, and is in charge of receiving the relevant information for the AR presentation on the mobile (3D models for plans, Object recognition models, etc.), detecting the location and visualizing the plans through the camera view.
C13	eGovernment Dialogue and Visualisation client	This mobile component is responsible for the visualisation of initiatives, questionnaires and results on the mobile.
C17	Webapplication for Augmented Reality Configuration	This web portal is used for adding AR objects and training object recognition models.

Table 60: Urban Planning Gordexola Mobile application prototype: Live+Gov components

The prototype used in the First Field Trial was a slightly improved version of the Urban Planning Gordexola mobile prototype described in D5.2 [1]. The actual version used for the trial was version 1.3.

This version of the mobile application for Android contained three hypothetical urban plans of Gordexola, representative of the plans that could take place in Gordexola along a political term. In this version for the first trial the plans were the ones described in D5.2: improvement of the town square, restoration of an old trail and the construction of a pedestrian bridge over the river. The application is totally in Spanish and has been slightly adjusted in wording. Moreover, the exact location and the images used to trigger the recognition option of the AR visualization have been somewhat fine-tuned. More detail about these changes will follow in the customized prototype evaluation section.
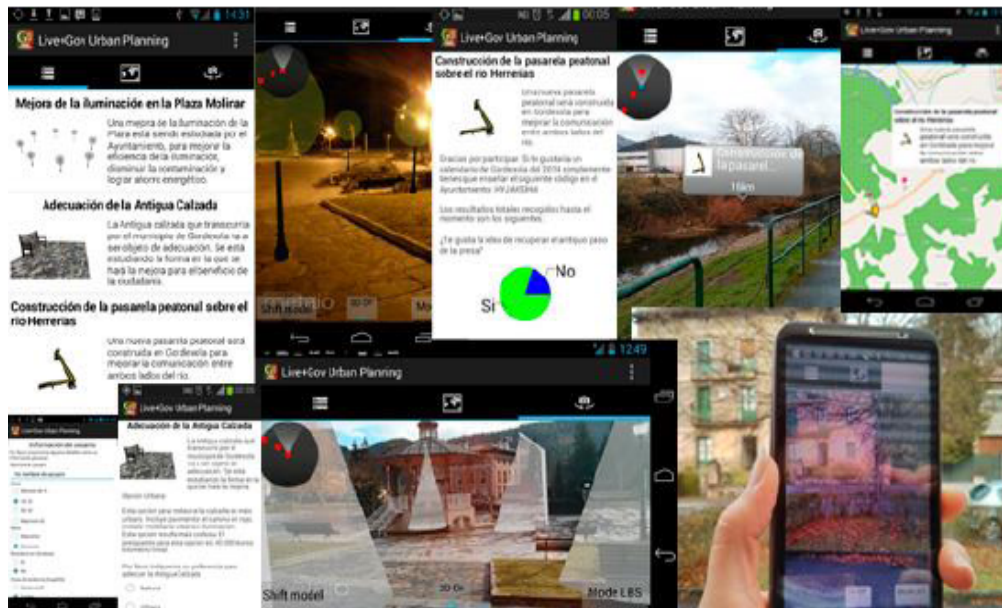
Figure 31: Impression of the prototype used for the 1st trial of Urban Planning

6.1.3 Evaluation process

Before the actual involvement of stakeholders in the work meetings, system testing in the real world scenario was conducted in Gordexola, once the prototype demonstrator was ready. Special attention was placed on this task to have a solution of enough quality to be shared with citizens due to the political responsibility that an initiative of this type can have for the administrators in a municipality, and therefore for the first trial to have a positive outcome.

As the whole application depends on location detection this was vital, as well as the presentation of the 3D models over the camera view in order to satisfy users. Therefore some extra fine-tuning and customization was done prior to the stakeholder presentations.

The first problem encountered when preparing the system testing was related to the devices. A set of Android Smart phones were ready for the trial once the application was delivered. However, it was impossible to use these devices as the application could not be installed in these devices. After studying the problem and trying different options, it was found that the reason for this was that the location detection and AR presentation required some hardware components of sufficient quality, such as those usually available in middle-end devices, in order to work, and therefore not all android smart phones were compatible even if they were recent. This device limitation, was not only an unfortunate coincidence of the selected trial devices, but also affected the devices that that the selected users had, and they were contacted prior to the meeting to make sure enough number of compatible devices would be available in order to conduct the participation experience. In other words, the trial users in some cases did not have own devices of sufficient quality, and had to borough high end devices from others for the trial. However we must stress that what a priori is a big drawback can be also considered a sign that the characteristics of the application developed are quite modern in this scenario, and we can predict that at the end of the project life more and more smart phones will comply with the requirements.

The application was tested in the planned locations with several devices from different phone companies at the same time. The results showed how the influence of mobile

location detection was not very accurate. This depends on several factors, which we had already analysed: mobile network, Wi-Fi hotspots, and GPS. This was considered as part of the limitations the system has, and that must be taken into account for future use, with special relevance here due to the fact that Gordexola is a rural area with a widespread mobile network. Even though 100% accuracy was not expected the behaviour of the visualization on different devices was quite different from one to another. Slight location misinterpretations could be handled for the Bridge and the Town Square light poles, but the initial location of the Renovation of the Pedestrian Trail was not feasible due to unforeseen geographical restrictions and a new location was selected. Therefore although the specification of devices required for this trial were high, and were not unforeseen, it still became an issue due to the previously described circumstances.

Some adjustments were also made to the images that were used to trigger the object recognition for the image based augmented reality visualization. Some frequency treatment of the pictures improved the detection of the images in order to avoid some problems with the models moving or flipping that were detected during system testing.

These reasons, together with the Christmas holidays, delayed the actual field trial meetings to the end of January.

Protoype evaluation				
Aspect	Criteria	Specifications	Targeted WP	Methods used
Customization process success	Technical: Integration	How well were the included components integrated?	WP4 & WP5	Internal review
	Technical: Functional	Were all given requirements met?	WP5	Internal review
	Non-technical customization	Did the application comply on field?	WP3, WP4 & WP5	Internal review
User activity	Quantity of data	Was enough data collected from users?	WP2 &WP5	Data analysis
	uantity of system data for analysis	Was enough data collected from the system?	WP3, WP4 &WP5	Data analysis
	Interest	Did users express their interest in using the application?	WP5	Questionnaires, personal contact
	Coverage	Did enough users take part?	WP2&WP5	Data analysis, questionnaires, personal contact
Usability	Learnability	How easy is it for users to accomplish basic tasks the first time they encounter the design?	WP5	questionnaires, personal contact
	Utility	Does it do what users need?	WP5	questionnaires, personal contact

		When users return to the design after a period of not using it, how easily can they reestablish proficiency?	WP5	questionnaires, personal contact
	Memorability			
	Satisfaction	How pleasant is it to use the system in general?	WP5	questionnaires, personal contact
	Efficiency	Once users have learned the design, how quickly can they perform tasks?	WP5	questionnaires, personal contact
	Feedback & Errors	How many errors do users make, how severe are these errors, and how easily can they recover from the errors?	WP5 & WP4	questionnaires, personal contact, data analysis
	Reliability	Does the software run consistently without crashing?	WP5 & WP4	questionnaires, personal contact, data analysis
Quality	Data	Is the quality of the information offered satisfying for citizens?	WP2 & WP5	Questionnaires, personal contact
		Is the quality of the information collected satisfying for administrators?	WP2 & WP5	Questionnaires, personal contact
	Augmented Reality	Does the enhanced view give added value to the participation initiative?	WP2 & WP5	Questionnaires, personal contact
Impact	General	User numbers	WP5	Data analysis
		Number of participations	WP5	Data analysis
	Authority	Does the application provide useful insights for decision making?	WP2&WP5	Personal contact, questionnaires, comparison to existing tools
	Citizen	Does this have an impact on citizens' participation in public life?	WP5	questionnaires, personal contact
	Interaction	Does the application provide a useful tool for interaction?	WP5	questionnaires, personal contact

Table 61: Evaluation aspects and methods in the 1st trial of Urban Planning Gordexola

6.2   User-experiences and customised prototype evaluation

6.2.1   Evaluation of Usability

The user experience of the use case will be evaluated through indicators that are in some cases countable and in others closer to feelings and impressions, as was stated in D5.1. [2]

For the first trial, as the citizen trial has been with a set group of selected users, the overall numbers are only a relative indicator. The user experience has been collected using a questionnaire following the guidelines mentioned above to study the learnability, utility, memorability, satisfaction, efficiency, feedback and errors and reliability. These questionnaires have been complimented by personal interviewing in order to better share impressions with the users.

The use of the system in the first trial can be quantified by the following indicators. These numbers are low, but must be interpreted as the result of a small-scale first trial, that can offer insights on what we can expect when the second trial opens to a greater number of users.

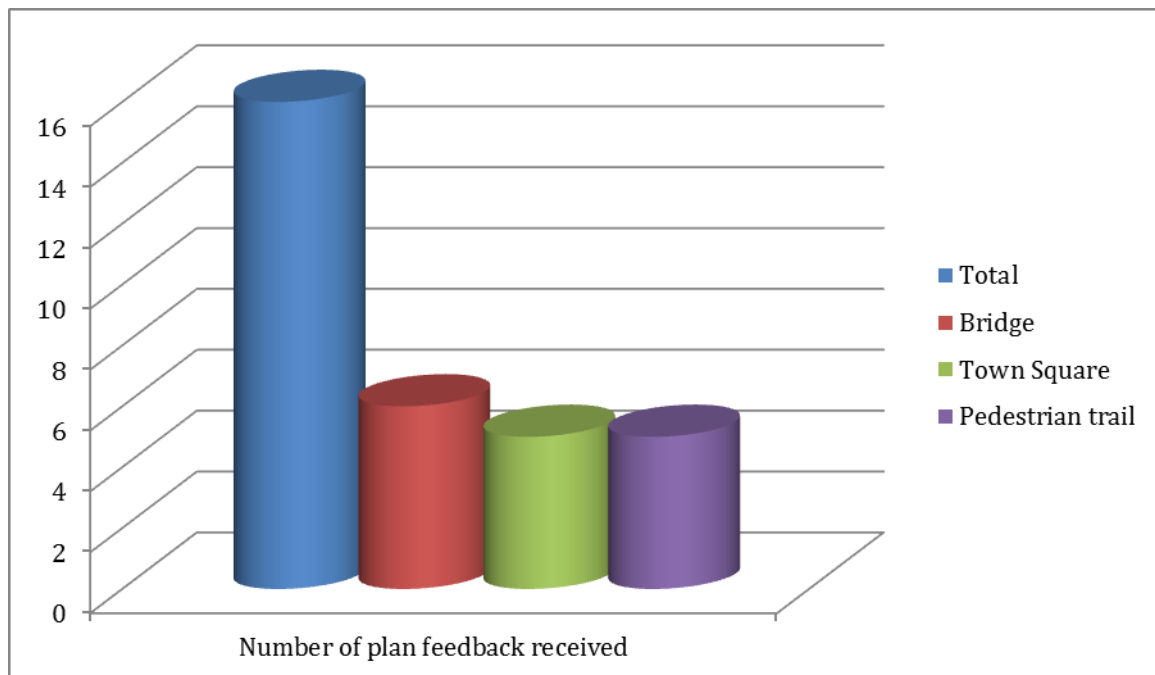Figure 32: Number of plan feedback received Urban Planning Gordexola 1st trial

Looking at general numbers, on average, users that were able to install the application participated by voting in 1.5 plans (Figure 32). However, looking closer at the results of the participation we can see that roughly 50% of the users were interested enough or able to participate in all of the options they were offered (Figure 33).
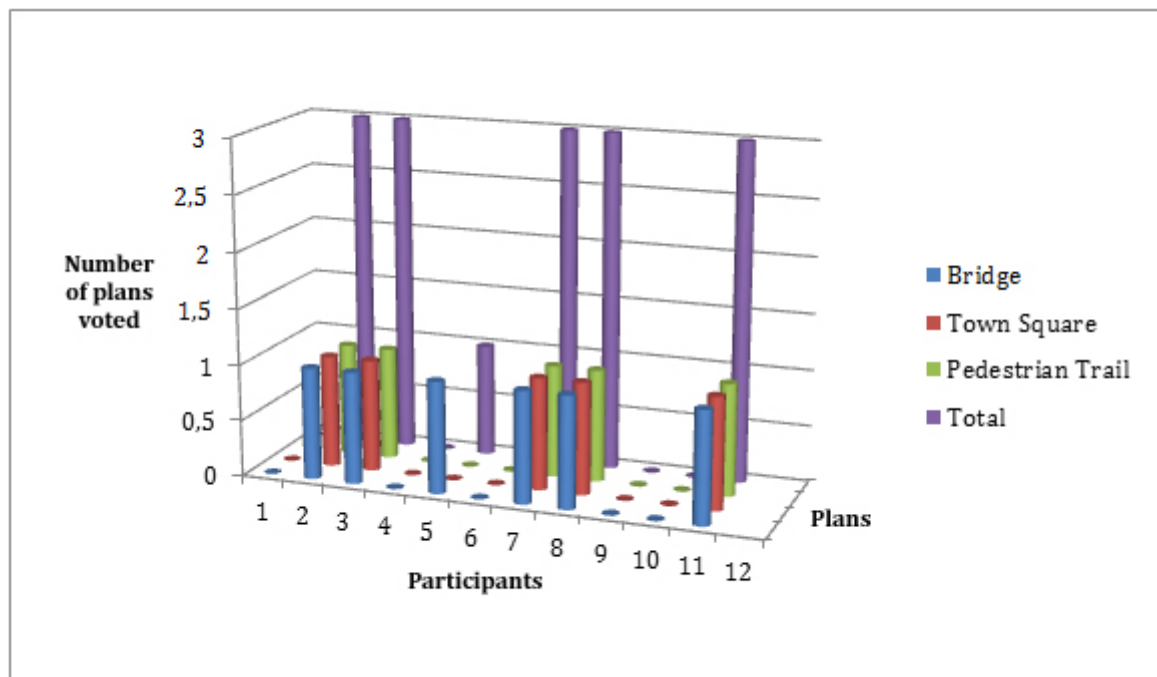
Figure 33: Number of plan feedback received from each participant Urban Planning Gordexola first trial

Another analysis that was planned for evaluation was comparison to other participation initiatives. As there have been no precedents of this type we were unable to use this indicator. Other indicators related to the total plans in the town or the number of plans affected, are not relevant indicators in this case as the plans are not real and it is not possible to compare in these terms with other plans.

Another aspect planned on being evaluated was the customization of the application. This process can be divided in the technical and non-technical aspects. The technical aspects have been mentioned before, and are also detailed in the specific evaluation for WPs. However the most challenging part has been the non-technical customization as many difficulties have arisen. Mainly location and 3D models have been problematic, but the evaluation of the experience is positive. In the first case there were unforeseen geographical restrictions already mentioned and in the second case 3D models were more difficult to get than initially expected.

The evaluation of citizen opinion in terms of the look and feel and the experience with the technical innovations have been handled through the previously mentioned questionnaire that combine questions for all of the aspects as well as comments given in the session, and the results are further detailed in the following sections. The raw results are presented in tables and the scale used was from 1 to 5 (1 = strongly disagree, 2= disagree, 3= neutral, 4= agree, 5= strongly agree).

Each section starts with the main question that is evaluated.

6.2.1.1 Learnability

How easy is it for users to accomplish basic tasks the first time they encounter the design? This was asked to citizen users through a set of questions in order to measure their feeling towards this aspect. The average results collected in the surveys that are in the following

table show that users were satisfied with the application in terms of clarity and that they considered it was intuitive.

Aspect	Question	Average
Learnability	I learned to use Urban Planning Gordexola quickly	3
	It is user friendly	3
	Participating is straight forward	4
	It is easy to explore new features by trial and error	3
	The wording is clear and it is easy to understand the message	3
	It is easy to find the information I needed	3
	I can use it without written instructions	3
	Overall, I am satisfied with the amount of time it took to participate using Urban Planning Gordexola	3
	Both occasional and regular users would like it	3
	The grouping of menu options are logical	3

Table 62: Learnability results from UPG questionnaire

Therefore learnability of the application has been given a positive grade. Although the results are positive, and for the second trial we shall try to keep this level, attention must be given to this aspect. In the first trial there was direct information and support during the session, which together with the user manual, helped guide the users. For the second trial there will be more sessions, and support will be given via telephone by BIZ.

6.2.1.2 Utility

Does it do what users need?

The second aspect measures how useful the trial users find the application that they were presented with. This aspect could be considered the most important of all because without it being useful, even if the system complied to all other aspects it would be meaningless.

Aspect	Question	Average
Utility	It does everything I would expect it to do	3
	It makes participating in my municipality easier	4
	AR presentation helps me better understand the plan	4
	This app helps me know more about municipal decisions in Gordexola	4
	This app helps me better understand how the Local Council decides about plans	3
	I receive interesting information through the UPG app	3
	AR lets me see how the plan will look like in the end	4

Table 63: Utility results from UPG questionnaire

The results of the questions related to the utility of the application received a high average, being mostly satisfactory for users. Of course these results are based on the aim of participating without evaluating if this participation is taken into account, which will be the target for the second trial.

In terms of utility, one improvement suggested both in the group brainstorming with citizens and during the meeting with the councillors was to include a link to related information on a related web page. This way, those who are interested could receive more information, either directly on the mobile phone, or later on a computer, thanks to the link that they can access through the application on their phone.

6.2.1.3 Memorability

When users return to the design after a period of not using it, how easily can they re-establish proficiency? The results of the user experience relating to continuity are quite interesting in this application as the aim would be to have this application running along time. It is intended to be a communication channel that is used when there is a plan or issue to be discussed by both ends, therefore using it after some time should be easy to attract participants once and again.

Aspect	Question	**Average**
Memorability	Sequence of screens is very clear	**4**
	The organization of information on screens is clear	**3**
	The information collected from the user is adequate.	**4**
	The information received is interesting to base my opinion.	**3**
	I don´t notice any inconsistencies as I use it.	**3**

Table 64: Memorability results from UPG questionnaire

Memorability is directly related to learnability. As the application is very simple for users, using it after some time was also considered to be easy by the trial users, as the results show, slightly higher than those of learnability. An interpretation of this result could be that users find it easier to use the application to vote for a plan once you have used the application before, even though time has gone by.

However we must point out that due to the setup of the trial meeting this evaluation is actually a prediction made by users based on the simplicity and ease of use, and that the only repetition that they have been able to do is vote on the three different plans. It is interesting to mention that the results of voting collected showed how one of the trial users who participated in the meeting sent information through the application the next day.

6.2.1.4 Satisfaction

How pleasant is it to use the system in general?

The satisfaction in this case is in terms of the use made by the citizen of the application and the results of its use.

Aspect	Question	Average
Satisfaction	Overall, I am satisfied with how easy it is to use UPG	3
	I feel comfortable using this app	3
	I like using the interface of Urban Planning Gordexola	3
	I would recommend it to a friend	3
	The Local Council is interested in knowing my opinion and my needs	4
	I feel satisfied participating in local initiatives	4

Table 65: Satisfaction results from UPG questionnaire

In general the satisfaction results are also positive, this was also acknowledged in the personal impressions that trial users shared during the session. They were satisfied with the look and feel of the application and how it worked, except for the reliability, which will be discussed in its own section. Another aspect that worried some of the trial users was the battery use of the application, as several devices died during the trial.

Some satisfaction aspects, not directly linked to this application that was being used, but that are part of the first trial are the problems with downloading and installing, and the unconformity with supported devices. Users did not answer the questionnaire negatively although they experienced many problems with the installation of the application, due to the fact that they received it by email or Bluetooth transfer instead of downloading it from the official channels. For the second trial the apps need to be offered through regular app stores.

The same consideration should be made to dissatisfaction expressed by them due to the limitation of devices, as the application was only available on Android devices for the first trial. Initially they had the same opinion with the application only working on some smart phones. However, they understood the explanation given to them that the requirements for visualization of 3D models through Augmented Reality triggered by location detection that is a milestone for this application, depended on components of enough quality, only featured in these devices. This was also supported with the assumption that as time goes by higher quality devices will be a standard.

6.2.1.5 Efficiency

Once users have learned the design, how quickly can they perform tasks?

Aspect	Question	Average
Efficiency	It requires the fewest steps possible to express my opinion	4
	The required user actions are consistent	3
	Overall, I am satisfied with the amount of time it took to complete the tasks in this scenario	3
	This app is a new channel to communicate with my Local Council	4
	This app allows me to frequently evaluate local initiatives	4

	It is an efficient way to know the opinion of people in times of crisis	4

Table 66: Efficiency results from UPG questionnaire

The feedback collected from the trial users relating to efficiency was high rated. The efficiency of the application received a positive mark, but especially well considered was the efficiency of the system as a new communication channel to enhance the relationship with the local administration in terms of effective results. Again these answers are half way between the impressions and the predictions of the system in a hypothetical future long-term use.

Especially relevant feedback was collected about this aspect from the decision makers from Gordexola, as they evaluated the possibilities that the system offers as very efficient in terms of participation. Compared to current participation it would be more demanding from their perspective but results could be used for their decision making, which has been almost impossible up to now with their limited traditional participation experiences.

6.2.1.6 Feedback & Errors

How many errors do users make, how severe are those error, and how easily can they recover from errors?

Aspect	Question	Average
Feedback & Errors	The UPG app shows the user what is possible and what is not	3
	I can recover from mistakes quickly and easily	3
	Error messages are non-disruptive/ informative	2

Table 67: Feedback & Errors results from UPG questionnaire

Evaluation of the feedback from the running of the application itself detected that there was a clear lack of error messages. Users clearly stated that more feedback from the system was required as they did not have a clear idea of why they were suffering failures, or slow running of the application, or even why they experienced crashes. This aspect is directly related to reliability, and due to the problems with the application the lack of error information was strongly missed. As was already mentioned in the user satisfaction analysis, this had a great influence on users lowering their appreciation of the whole system due to uncertainty.

Feedback was also missed on data transfer and the information saved by the system, this is, what option the user finally selected.

6.2.1.7 Reliability

Does the software run consistently without crashing?

| Aspect	Question	Average
Reliability	I can use it successfully every time	2
	The 3D model is positioned correctly in the real view (location)	3

	The 3D model is positioned correctly in the image recognition based AR view	3

Table 68: Reliability results from UPG questionnaire

Although the results are not bad at first sight, if we analyse the global results and compare the answers to these questions individually a 2 on average was the lowest score. This once again shows that users in general did not use the lowest score, although there were some individual minimum scores that emphasized on this. Indeed, in relative terms reliability was considered the lowest aspect in the questionnaires.

This result was also collected in the interviewing during the session, as users shared their unhappiness with how the application worked in terms of crashing and that they felt uncomfortable as they were unable to use it. They said that they would abandon the use of this app with this level of reliability in a real world scenario.

Therefore the evaluation in terms of reliability is quite negative and needs to be evaluated by the technical work packages to considerably improve this aspect for the second field trial. This is crucial looking forward to a successful second trial where the scope of participants is planned to be wider and the commitment of the governing body also depends on the quality of the proposed system.
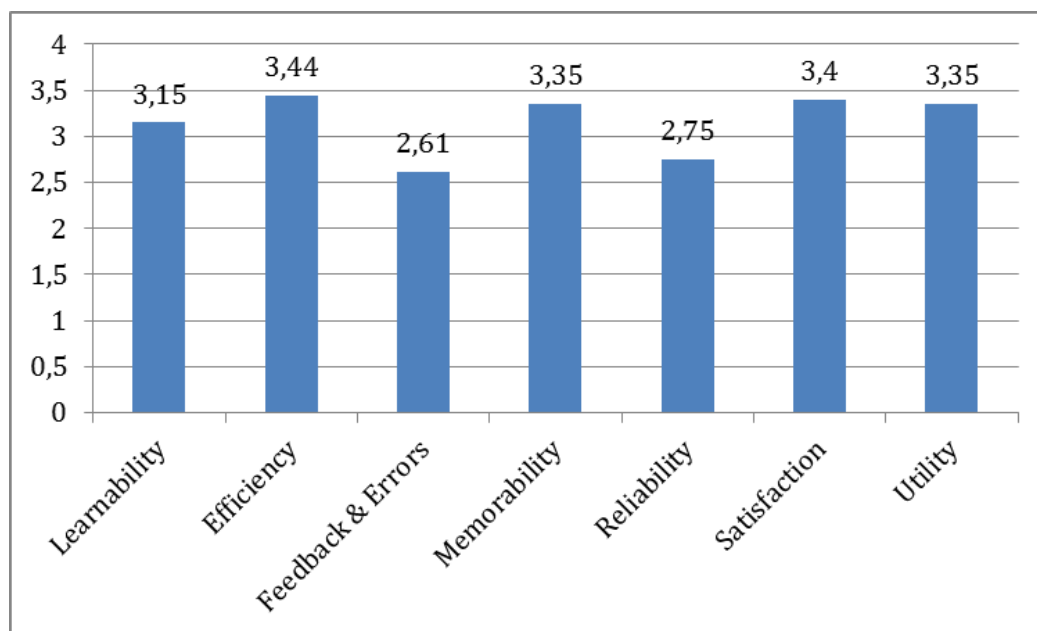
Figure 34: Average score on usability topics for *Urban Planning Gordexola*

In sum, a quick overview of the average scores of the usability topics that are presented in Figure 34 shows how the two topics that are below the score of 3, already identified earlier as the weakest points, are feedback & errors, and reliability. These numbers clearly address the issues encountered during the trial session that have been described.

6.2.2 Evaluation of the requirements

The starting point of the evaluation was to analyse if the initial requirements complied in D5.1 [2] were properly addressed in this first trial prototype.

Regarding the functional requirements for the mobile application, most have been fulfilled for the first trial.

L+G-FR.02	Be able to sense the context of users (e.g. location, means of transportation, activities etc)
Implementation	Location is detected in the application to show location based AR view and map view User context is also "sensed" by object recognition in image-based detection AR view.
Evaluation and refinement	This requirement is met in the first field trial for citizen trial users and is still valid for the second. Context of users will be further studied in WP1 to see how activity recognition can be useful for the use case. Refinement of location detection needs to be handled so that the map view and the location based augmented reality have the same values.

L+G-FR.03	Be able to show information in a context aware manner (location, person, time-specific, situational) to create greater awareness and create mutual citizen-government understanding
Implementation	Location based AR view and map view User context also triggers the image based AR view when there is object recognition
Evaluation and refinement	This requirement is met in the first field trial for citizen trial users and is still valid for the second. Refinement: geographical restrictions need to be contemplated and taken into account. Tags need to be revised to make their function straightforward for users.

L+G-FR.05	Provide communicative tools for government and citizens (including reactions, alerts, messages)
Implementation	Information flow from local administration to citizens: plans are presented, relevant information is included, 3D models offer a realistic visualization. The results of the real time total results of the voting for each plan are presented to the users after voting. Information flow from citizens to local administration: votes or opinion are shared and free text can also be returned for each plan or issue.
Evaluation and refinement	This requirement is met in the first trial. Refinement: more feedback about how the results will be used should be included.

L+G-FR.07	Be able to capture images and recognize their content
Implementation	This requirement is present in Urban Planning Gordexola in the image based triggering of AR content.
Evaluation and refinement	This requirement has been met in the first trial. Refinement: further work needs to be done so that the presentation of the models through image recognition is more precise. No moving of the 3D models superimposed on the triggering image needs to be achieved.

L+G-FR.08	Provide possibilities for textual feedback. For example how his contribution is taken into consideration, including possible benefits for participating and status report on the situation
Implementation	Feedback is provided to users about the total results of voting after casting the vote.
Evaluation and refinement	This requirement is met in the first trial. Refinement: administrators should be allowed to give textual feedback to users for the second trial.

L+G-FR.09	Allow a user to view all of the issues that are currently open, and also a history record of "my issues" (where the citizen has participated) and "my municipality" (where all of the past issues from the town can be found) with a short summary of each of the past issues.
Implementation	The list of currently open issues is presented in the list view.
Evaluation and refinement	This requirement is partially met in the first trial. Refinement: show a history record with past issues and somehow distinguish between those that a user has participated in and those that not.

L+G-FR.11	Provide (mobile) presentation possibilities in an Augmented manner
Implementation	Different AR views.
Evaluation and refinement	This requirement is met in the first trial and valid for the second trial.

L+G-FR.12	Provide possibilities to retrieve personal opinions, votes or comments from users
Implementation	Voting option and free comment box to collect feedback from citizens in the plan description screen within the application
Evaluation and refinement	This requirement is met in the first trial and valid for the second trial.

L+G-FR.21	Provide possibilities to communicate urban planning plans
Implementation	The application itself is a way of communicating the plans with users.
Evaluation and refinement	This requirement is met in the first trial and valid for the second trial.

However, the following have been postponed for the second phase, as they required input from the first trial. This is the case of offering feedback to the citizen user about how the plan is taken into consideration, included as a refinement in the previous requirement list (L+G-FR.05), as well as the functionality of offering plan history records (L+G-FR.09). The latest is being worked on in order to find the correct presentation of past participation in the best way from the point of view of enhancing the communication between citizens and administration. This will be worked on by the "working group". This is also the case for the filtering of issues, visualization based on knowledge and patterns from data-mining, using citizens as sensors and connection with administrative systems as they requires more analysis in order to see if this requirement still stands or if it is finally discarded.

L+G-FR.04	Provide possibilities to group, aggregate, filter large amounts of data to discover knowledge
Implementation	-
Evaluation and refinement	This requirement has been postponed for the second trial. It will compliment L+G-FR.02 activity recognition.

L+G-FR.06	Provide possibilities to visualize knowledge and patterns from data by data-mining techniques
Implementation	-
Evaluation and refinement	This requirement has been postponed for the second trial. It will compliment L+G-FR.02 activity recognition.

L+G-FR.10	Be able to utilize 'citizens as sensors' for their environment by mining relevant data (e.g. location, context, pictures, reports) from their reality
Implementation	-
Evaluation and refinement	Not present in Urban Planning Gordexola. Postponed for the second trial as part of the application for administrators. Refinement: extract location information of where citizens participate, the context of voting (L+G-FR.02)

L+G-FR.13	Be able to connect to different municipality administration systems
Implementation	-
Evaluation and refinement	Not present in Urban Planning Gordexola. Postponed for the second trial as part of the application for administrators.

The functional requirements of the reporting authorities have been taken into account in hand made reporting for the first trial. However, an automated system to present the collected information is planned for the second phase of the use case work.

6.2.2.1 Citizen

As explained above the trial of the Urban Planning Gordexola Prototype application was open to a limited group of 12 users. The overall evaluation of the trial from the citizen experience was that the prototype needed to be improved in order to be ready for the public trial as at this point it did not satisfy users and they were not able to use the application the way that was expected.

Figure 35: Citizens using image recognition during the 1st trial of Urban Planning

The results from the evaluation in general lines show satisfaction with the initiative and the idea of enhanced views of plans, which are many times complicated to understand in the current way of receiving information. However there were technical problems with the application that eclipsed the general user experience. Lessons learned were that the application needed to have contemplated error or warning messages to guide users. In numbers, one of the 12 participants was unable to install the app, and 5 participants voted for all of the plans. These results were due to technical problems and not lack of interest. The causes of the technical problems were difficult to identify during the session, and in some cases the complete application crashed.

The surveys that were prepared followed the general guidelines of usability questionnaires from WP5, where some use case specific questions were added. (see also Table 1)

6.2.2.2 Government

The Local Council members of Gordexola that participated in the First Trial were asked in this stage about the main issues related to local participation in Gordexola in general terms, and

to evaluate the three fundamental issues that are questioned when planning a participatory process applied to the Urban Planning Field Trial.

In first place the cost of opening a participation process was discussed. The council members initially addressed that a participatory process such as Urban Planning Gordexola, would be costly for the Town Council. Then they were asked to make a comparison with other participatory processes and their costs. The answer to this was that Gordexola, similarly to other municipalities in the region, were lacking participatory processes, that on another hand were being demanded by people, and were set on their agenda as a to do in the near future. In the case of Gordexola, the few participatory processes that have been held have not been costly in terms of financial costs, as they were conducted by members and personnel of the local council, but did have a high personal effort and a low or close to no effective results. Finally after the discussion, they evaluated that the cost of a system such as Urban Planning Gordexola could be worthwhile as it would give interesting information to back the political decisions they make and in some cases it could even save costs if plans that were presented were rejected by citizens. For citizens the savings in time and effort would be quite relevant as they could participate whenever they want, not having a fixed schedule such as with meetings, and this ease of participation would also be a way of attracting more people who wouldn´t participate in traditional participation dynamics due to these reasons.

The next issue questioned was the problem of complexity. In this case, the administrators were happy with the way plans were presented in a simple manner, easy to understand, and with meaningful information to allow citizens to base their opinion or vote on. However, the problem on complexity that they found when evaluating the first trial was related to the difficulty that the smartphones themselves have for certain part of citizens. Therefore, although it is an interesting solution that looks towards the future, the evaluation of this aspect determined that the phone based participation needed to be backed by other complimentary participation means in order to not attend only the needs of a certain part of population and to avoid a digital gap when opening the participation process.

The third fundamental issue the administrators were asked to evaluate the trial on was the representativeness. This was discussed in the previous points already. In terms of evaluating the first field trials representativeness, a favourable aspect highlighted by the administrators was that the participation results had valuable information in order to segment the results and study the representativeness in terms of age, gender or area of residence, which they considered quite relevant in frequent medium-scale urban plans.

In general terms of representativeness the first field trial could only show the potential that could be reached in the second trial. Nonetheless, another result of the evaluation of representativeness in this first stage was the necessity to establish a threshold in participants for the way the results will be taken into account to avoid questions about the legitimacy, as well as a the plan for how to handle extraordinarily high numbers in participation that would require further actions. The administrators said that if a plan presented through this system showed that it was subject of an important public debate, it would be handled in an appropriate manner with public meetings, or in other ways depending on the case. This would also help in the case of organised groups of interest that back or protest against a plan and try to actively gather people to vote for or against a plan, trying to direct the results. However, they see this risk low as they would easily know that something like this is happening in the municipality.

6.2.3 Evaluation results for specific WPs

6.2.3.1 WP1 – Reality Sensing and Mining

Little can be evaluated from the first trial of Urban Planning for the work in Reality Sensing and Mining as no results of this work package have been included in this trial.

However, the general evaluation of the first trial of Urban Planning Gordexola can be used by WP1 to look into the possibilities that the scenario of this use case and its first field trial offer to extract useful information to improve this technologically advanced participation initiative. This will be particularly important to enhance the information that will be given to decision makers in the administrative part, and attention will be paid to this in the Working Group meetings with members of the Governing Body of Gordexola.

6.2.3.2 WP2 – Policy Modelling and Visualization for Mobile-aware Dialogue

The urban planning Use Case is focused on citizen participation. Therefore an analysis of how the advantages of public participation have been identified in the first field trial could be considered the starting point for its general evaluation, but also a result of evaluation for WP2. These advantages are included in D2.1. [5]

The First Trial of the Urban Planning Use Case pays special attention to the general concept of mobile citizen participation, as the engagement of the administrations was the main goal of this trial. Citizens are demanding more participation nowadays, also in the town of Gordexola; however what is challenging in this Use Case is that the opinion collected through a mobile application is taken into account in the local decision-making, and this depends on the confidence of administrators and their commitment.

The first advantage described is proactive adjustment of policies to the needs of citizens, taking into account the demographic groups in society. During the first trial these aspects have been identified in the Use Case as the application collects information about age, gender or residence. This allows administrators to know the needs of the citizens they represent together with demographic characteristics of those who give their opinion for the plans they are invited to participate in. This way further analysis is allowed relevant for the plan or decision when it is the case.

The second benefit stated was the access to the knowledge of the society. During the trial, the administrators themselves pointed this out as an advantage that they clearly saw the system offers them. It allows them to know what people think: the opinion of people that live in an area where a plan is relevant or people who have reasons to be pro or against a plan. Moreover if the participation is extraordinarily high, this tool lets them know that the plan is considered important for their neighbours.

The third advantage was related to early knowledge of important problems. This was also brought up by administrators, who saw how this tool would allow them to give people the option to decide upon an issue before starting the real development, resulting in significant savings for the town if a plan is decided not to be of interest for the citizens.

The fourth improvement would be related to communication between people and administration. This participation tool also gives information to citizens, and collects their feedback, in a manner that it can be compiled. This by itself increases trust by citizens that know what is happening. Moreover, administrators identified that communication with their citizens would also improve thanks to the way of presenting models on maps that are more user friendly and 3D visualization on the mobile device. They said that the standard way of

sharing information related with plans with the people who are interested is by showing the blue prints to those people who go to the Town Hall in person. This is normally a problem as many people don´t understand the blue prints, and they need a lot of extra explanation by the workers. The urban planning system with the location based view, and the 3D model presentation, gives a better understanding to people.

The fifth advantage would be attracting people who are normally not interested in politics. An easy participation system on the smartphone such as the application developed for the Urban Planning Use Case has this appeal, especially for young people that currently don´t share their insights with the Municipality.

The sixth benefit is reducing the influence of particularistic interest and lobby groups, and comes from a better knowledge of general opinion. This was also addressed by the administrators who were presented with the system as part of the field trial. This benefit is backed by the system only allowing people to vote once from their smart phone.

The seventh advantage was that this type of participation can raise awareness of things that are not directly visible. This is strongly related to the aspects mentioned before about having a greater knowledge coming from a larger number of people, and the administrators of Gordexola also identified this as a very interesting outcome of the trial. This system can also trigger other types of participation about issues that greatly concern the public opinion that they might have not identified through other means.

The eighth advantage is that opening participation will help more people have better insights about how decisions are made. This is also considered as a great outcome of the system by the Councillors of Gordexola.

Therefore, in general lines, the theoretical considerations from D2.1 [5] have been largely confirmed in the 1st trial of Urban Planning

After going through the general aspects, we could say that the main result of the first trial that is relevant as output from the trial and input for WP2, is the greater than expected commitment by the Governing Body of Gordexola. They positively received the results of the first trial although they required a higher level of quality of the citizen app for the second field trial.

More specifically the evaluation of the first trial has detected the need, expressed by both citizens and administrators, to develop a plan to reach people who are not smartphone users and bridge the "smart phone gap" and collect a broader collective opinion, not only of the segment of "modern smartphone owners". Several of the previous points also address positive aspects of finding a solution to this problem through the use of this system. Therefore, this is a weak point to attack the use case and that has been a result of the evaluation of both groups of stakeholders that however can be turned around if it is considered as the enhanced part that looks into the future, of a participatory process that includes all of the citizens that are interested in participating. The mobile phone cannot only give easy access for those who use one, but it can also be a tool for current participation. Most people visit the Town Hall in person to receive information about plans and give their opinion but there is no standard manner in which it is dealt with.

6.2.3.3   WP3 – Mobile Augmented Reality

The evaluation of WP3 through the first field trial of Urban Planning has verified that AR is quite resource demanding for the application and therefore for the smartphone, as well as

for the internet connection. This was not unforeseen but turned out to be an issue in the real world experimentation.

The way the 3D models are downloaded, as mentioned before, required a good internet connection and caused problems in the scenario of the trial. Also, the storage in files was considered inconvenient for users as they are difficult to delete, which is even more problematic due to the big size of the models.

On another hand, further evaluation of the influence of location will be needed to better understand the geographical restrictions of the system. Furthermore, adjustment is also needed between location used in the AR view and the map view for coherency.

The quality of 3D models will also need focus on, which is planned to be revised in collaboration with the local entity in the Working Group sessions, in order to improve the information given to the citizens and the experience using the enhanced camera view using the Augmented Reality functionality.

Revision of the labels and tags used for the AR view needs to be done so that they are self-explaining for the user.

As for the revised requirements L+G-FR.03, L+G-FR.05 and L+G-FR.11 are directly related to this work package and need to be taken into account. They are either met and valid for the second field trial in the current implementation or need to be refined as has already been analysed in section 6.2.1.

All of these factors need to be taken into account to further detail the AR system needs and reach a technically feasible system on the level needed for the scenario.

6.2.3.4   WP4 – System Integration

The components used in the Urban Planning Trial application are successfully integrated. Multiple components from the Live+Gov Toolkit are utilized. In this trial the Service Center, Augmented Reality and eGoverment and Dialogue components are combined to create the customized solution for Urban Planning. During the trial the components proved to be able to work together correctly in an integrated solution.

The high demand of resources by the AR components that has been detected must be taken into account as a result of evaluation from the point of view of integration due to the general consequences one component has on the whole system in order to reach a technically feasible system appropriate for the scenario. The experience with these problems gave evidence that there is need for more status information on e.g. messages warnings provided to the user. Even though the application was tested in detail by project members before the launch with positive results, it appears that tests done with more variety of devices and android versions could have helped to better detect problems before or during the trial. It is also clear that due to short development cycles, the full needed attention was not possible to give in the integration and testing thoroughly and with the variety that would have been required to provide fully reliable application.

The decision of distributing the app via e-mail, taken by BIZ, turned out to be a mistake. The option of distributing via Google beta is a lesson learned for the second trial.

As for the revised requirements, special attention should be given from the perspective of WP4 to L+G-FR.01, L+G-FR.05, L+G-FR.08, L+G-FR.09, L+G-FR.12, L+G-FR.21, and L+G-FR.13 in section 6.2.1.

6.2.4 Impact and success of the Use Case trial

The impact of the first trial of the Urban Plan is quite limited in numbers due to the conception and political relevance it has. Opening a participatory process of this type at the local level in Biscay requires a lot of political commitment from the involved local administration as there are no similar initiatives and it changes the actual conception of representation in the local political arena. Although we are working in a project with prototypes, it can have negative political consequences if it is not a success.

Therefore we will study the impact of the trial and the use case in terms of the commitment that the Governing Body of Gordexola has made towards the second field trial. This impact is quite big as this group of representatives has agreed on opening a real plan for public participation for the second field trial, find an appropriate plan to present to their community, and introduce the result in the decision making process. The way in which the results will be incorporated in the current decision making process or the weight they will have in the decision will also be determined during future periodical "Working Group" meetings which will be conducted looking forward to the second field trial. They have also agreed to work on acquiring real 3D models of acceptable quality as part of their collaboration that they will try to get relating to the final urban plan considered.

Their acceptance of the project goals is very motivating at this point of the project, and this can definitely be considered as internal impact in the project, as they are ready to go one step forward and make the participation real. This will definitely have reflection externally.

The Councillors also agreed on studying the best way to do dissemination of the initiative, including their own media (website, magazine and printed material) and press releases, to impulse the public projection of Urban Planning and of the Live+Gov project.

6.3 Summary of the evaluation of Urban Planning Gordexola

The first field trial was conducted as a double session of meetings with citizens and administrators. The citizen trial was conducted to introduce the use case, and more specifically the application to a selected group of trial users and afterwards collect their impressions. Later on a meeting with the administrators of Gordexola was held, to do the same as an introduction, and then to present the results that would be input for their current decision making process, also collecting their impressions at the end.

The overall summary of the evaluation of the first field trial of Urban Planning Gordexola is that the key goal of the trial has been met as commitment from the local administrators has been achieved and interest has been raised among the citizen trial users who have been introduced in the initiative. However, this initial predisposition will need to be backed by the second field trial where the result of the commitment will be measurable.

Other results of the first field trial in Gordexola have been that although the application worked correctly during the testing, it wasn´t completely stable, presumably due to the requirements of the AR component. This higher demand than expected in terms of smartphone and connectivity needs became visible only during the trial. The system testing was conducted and there was an acceptance test prior to the trial. Initially, unforeseen geographical restrictions were experimented and minor adjustments were made on the application. However real world scenarios proved the testing was insufficient and also showed that there was a lack of error feedback to the citizen users.

Due to technical and organisational reasons the first trial was not as smooth as it was intended to be, as has already been described, and problems appeared during the trial and not before. Towards the second cycle, the lessons learned will translate in greater efforts in order to do a more exhaustive on field testing and not underestimate the problems that can appear. The level of preciseness required for the Augmented Reality functionality is high and this high level technical approach is the reason for conducting field trials. Therefore the results, although not totally positive, help working on the technical feasibility of the system on this scenario.

The evaluation of usability in terms of learnability, utility, memorability, satisfaction, efficiency, feedback & errors and reliability, has been positive by users in general lines, although improvement in reliability and stability has been highlighted, as well as in feedback & errors. The evaluation of the system by administrators has also been quite positive at this point with the information they have received as part of first trial, and they are engaged with the work for the second round and optimistically awaiting the results of the project and seeing how participation can be improved at their municipal administration level.

In sum, many insights have been acquired during the trial and the main aim of engagement that was set for the first trial has been reached. Therefore we can consider that the first trial has been successful in its goals, although work will be done to notably improve for the second cycle.

7 Summary evaluation results of the Functional Requirements

With our three use Case trials, consisting of four scenarios, a tremendous amount of information for upcoming period of Live+Gov has been collected. This information is on a large variety of aspects: usability of our customised prototypes, requirement revisions and user experiences, input for specific WPs. Apart from this we generated stakeholder engagement, media attention, political and organisation adaption. Furthermore, we have gained detailed insights on our SaaS component-based software architecture and the stability of our technical components in field settings. The field trials provided a large amount of information and insights on the organisational and political domain of eParticipation and the transition process (local) governments are dealing with. Finally, lessons are learned for the execution of the next field trial cycle, which will be of high importance for the Training Package and Overall Live+Gov methodology.

In this concluding chapter, main revisions and results of the Use Case evaluations for the implemented requirements of the first field trials are presented. These results are of importance for the second development, integration and customisation cycle. Most detailed results can be found in the specific trial sections 3-6. In Table 69, the implemented requirements as described in D5.1 [2] are provided with the specific Use Case evaluation results of the first field trials. These results and revisions will guide the upcoming period for Live+Gov.

Nr	Live+Gov system functional requirement	Mobility	Urban Maintenance	Urban Planning
L+G-FR.02	be able to sense the context of users (e.g. location, means of transportation, activities etc)		L+G-FR.02 has been met successfully; the quality of the Issue reports is regarded as very high. For short-term cycles, the information provided give municipality great insights on specific Urban Maintenance issues. On the more abstract level of citizen experiences, the mobile client provides municipalities with a large amount of information of the experienced state of the life-worlds. The requirement links to L+G-FR.10 and L+G-FR.19 as well. [Issue Reporting]	This requirement is met in the first field trial for citizen trial users and is still valid for the second. Context of users will be further studied in WP1 to see how activity recognition can be useful for the use case. Refinement of location detection needs to be handled so that the map view and the location based augmented reality have the same values.
L+G-FR.03	be able to show information in a context aware manner (location, person, time-specific, situational) to create greater awareness and create mutual citizen-government understanding		L+G-FR.03 is met successfully. Next steps of investigations and development need to focus on inclusion of mobile device support (tablets in particular). Furthermore, *Jij Maakt Utrecht* can benefit of clear explanations of the type of information that is presented. A satellite option for the map-view will be considered for future development. Lastly, concrete linkage to social networks like Facebook need to be considered. [*Jij Maakt Utrecht*]	This requirement is met in the first field trial for citizen trial users and is still valid for the second. Refinement: geographical restrictions need to be contemplated and taken into account. Tags need to be revised to make their function straightforward for users.
L+G-FR.04	provide possibilities to group, aggregate, filter large amounts of data to discover knowledge		L+G-FR.04 is met positively by the users. The webapplication is technically able to handle a large amount of data, which is one of the biggest challenges. The possibilities of discovery of knowledge is currently on a basic level, but further advancements on data-mining and knowledge discovery are planned for upcoming period in WP1 (see the link with L+G-FR.06). [*Jij Maakt Utrecht*]	

L+G-FR.05	provide communicative tools for government and citizens (including reactions, alerts, messages)	This requirement was met with the following features: issue reporting jam detection alerts exception information messages authority messaging Multiple communication channels were implemented in the prototype. These features include communication between citizens and authorities in both ways. Therefore we can state this requirement to have been met successfully.	The implementation of the requirement L+G-FR.05 as assessed initially is experienced positively. However, during the co-creation process and trial, new ideas and new insights on the aims for the requirement provide us with possible enrichments and specifications on the aspect of communicative tools. Concrete deep-link options need to be investigated for future development as well as forms of active notifications for new reactions and initiatives. [*Jij Maakt Utrecht*] L+G-FR.05 has been met with great success. 14.291 reports were accompanied of textual feedback by our 14 trial municipalities. The implementation on our running production system went smooth, even with the support of 6 different platforms (iPhone, Android, BlackBerry, Symbian, Windows Phone, and Windows Mobile). Future exploration will be done in possibilities of photograph feedback and notifier options. [Issue Reporting]	This requirement is met in the first trial. Refinement: more feedback about how the results will be used should be included.
L+G-FR.06	provide possibilities to visualize knowledge and patterns from data by data-mining techniques		The requirement is met successfully on the aspect of the visualisation. The content and patterns presented are currently not state-of-the-art yet, however this is conform the Live+Gov planning. Further developments in upcoming period will create the possibilities to present real knowledge patterns. L+G-FR.06 is currently at a stage that is proves itself as a firm foundation for future developments. [*Jij Maakt Utrecht*]	

L+G-FR.07	be able to capture images and recognize their content	This requirement was met with the issue reporting feature, where users could attach images to their issue report. AR has not been included in this trial and therefore the content recognition has not been implemented in this feature, but the requirement in terms of the use case has been successfully met.		This requirement has been met in the first trial. Refinement: further work needs to be done so that the presentation of the models through image recognition is more precise. No moving of the 3D models superimposed on the triggering image needs to be achieved.
L+G-FR.08	provide possibilities for textual feedback. For example how his contribution is taken into consideration, including possible benefits for participating and status report on the situation		More than 70% of the reports in our trial have been provided with textual feedback (14.291 reports in total). L+G-FR.08 is met highly successful. The option has been praised by both citizens as well as public officials. For next steps we will consider notifiers (push mechanism could be a possibility) and the possibility for municipalities to provide the reporter with a photograph (of the situation after it has been dealt with). [Issue Reporting]	This requirement is met in the first trial. Refinement: administrators should be allowed to give textual feedback to users for the second trial.
L+G-FR.09	allow a user to view all of the issues that are currently open, and also a history record of "my issues" (where the citizen has participated) and "my municipality" (where all of the past issues from the town can be found) with a short summary of each of the past issues.		L+G-FR.09 is met successfully: the user is provided with a clear overview of his reports, with possibilities to look for further details for each of them. [Issue Reporting]	This requirement is partially met in the first trial. Refinement: show a history record with past issues and somehow distinguish between those that a user has participated in and those that not.

L+G-FR.10	be able to utilize 'citizens as sensors' for their environment by mining relevant data (e.g. location, context, pictures, reports) from their reality	This requirement was targeted by collecting journey data and issue reports from citizens. From both, more advanced information is extracted with data analysis and visualization methods and therefore this requirement has been met successfully.	L+G-FR.10 is one of the crucial requirements of the Urban Maintenance Use Case. Within the trial close to 20000 detailed reports have been made, of which 15357 came with a photograph (77%). All reports have the exact location and around 85% are accompanied of a user description. These reports, result of citizens as sensor techniques, give tremendous information to municipalities. L+G-FR.10 is successfully met. [Issue Reporting]	Refinement: extract location information of where citizens participate, the context of voting (L+G-FR.02)
L+G-FR.11	provide (mobile) presentation possibilities in an Augmented manner			This requirement is met in the first trial and valid for the second trial.
L+G-FR.12	provide possibilities to retrieve personal opinions, votes or comments from users			This requirement is met in the first trial and valid for the second trial.

L+G-FR.13	be able to connect to different municipality administration systems	Targeted for this requirement, the prototype included connection to the real-time information system from where information was collected and utilized in service line detection, and traffic jam detection. Also, the disruption info messages were obtained by communicating with an external system maintained by HSL and shown via the application to the users. With two different systems successfully connected, we can state also this requirement to be met in the trial.	Interesting insights have been gained from the trial. Contrary to initial assessment, data interesting for Jij Maakt Utrecht is not available in any format that fits a true external connection. For the trial, we overcame this by organising data-collection sessions with the customised data-collection prototype. However, at this point we assess L+G-FR.13 as having low priority for *Jij Maakt Utrecht*. This means that we discontinue L+G-FR.13 for Jij Maakt Utrecht, because of practical reasons. We do want to note here that we are confident that our architecture is perfectly able to connect to external administrative systems. [*Jij Maakt Utrecht*] Within our trial we have proven with the StUF connection of the Issue Reporting client with the backend system of Eindhoven that we're able to interact with external administrative systems. L+G-FR.13 has been successfully met in our trial. This success is especially relevant for WP4 in which the possibilities to connect with public systems is located. [Issue Reporting]	
L+G-FR.14	receive data from the application containing travel information and means of travel	The prototype collects movement and journey data recorded by the users based on location and time. Also specific components for detecting activity (HAR) and used service line (SLD) were implemented. With these features, the prototype contains several points that together meet the given requirement. Thus, this requirement is met in the first trial.		

L+G-FR.17	detect traffic jams	Based on this requirement, a specific traffic jam detection module was developed and implemented. During the trial, warnings of jams have been generated by the module and the system has processed warnings to the users. With this, also the final requirement has been met.	
L+G-FR.19	collect and mine citizen experience and citizen (implicit) norms		The conceptual change as experimented with for the City Safari has been interesting. Whereas with our normal issue reporting, the collection of citizen experiences is focussed on 'issues', in the City Safari people were able to send 'reports of praise' as well. For future steps, we see opportunities to explore the possibilities of positive ratings.

Linked to the mining experiences and implicit norms, we want to explore the possibility to communicate existing service level norms. This would inform the citizen greatly, and would make it possible to match the experienced norm and the official quality level.

With these possible future steps in mind, we conclude that L+G-FR.19 is met successfully. [Issue Reporting] |
| L+G-FR.20 | provide possibilities to present co-maintenance initiatives in an area | | L+G-FR.20 has been fully met. Both citizens and authorities praise the completeness of information that wasn't available at all before *Jij Maakt Utrecht*. The combination of the gallery view and the map-based view is regarded as both informative as well as visually attractive. [*Jij Maakt Utrecht*] |

L+G-FR.21	provide possibilities to communicate urban planning plans			This requirement is met in the first trial and valid for the second trial.
L+G-FR.25	improve the quality of existing (open) data sets related to urban space, by leverage of citizens functioning as sensors and verify with existing registrations		For the scope of the Urban Maintenance trial, *Jij Maakt Utrecht* is a catalyser for collecting data about co-maintenance. New insight for the municipality was that this data was not available. We can conclude that we met L+G-FR.25 beyond success as we gathered the data with the citizen-as-sensor approach and not just improved the quality of existing data. Apart from data within our scope, *Jij Maakt Utrecht* holds a diverse range of data-sets. The ambition to improve the quality of all these sets is too ambitious. On the level of transparency, the publication is successful. However, the municipality of Utrecht concludes that we should be careful with bringing too broad goals into the application. As Use Case owner, we agree with the strategy that it is better not to expand to too many different domains within the webapplication. [*Jij Maakt Utrecht*]	
L+G-FR.26	introduce elements of serious gaming / gamification		The concept of gamification has been explored and an experiment has been executed. Although perceived as interesting, the requirement is regarded as too ambitious by the municipality of Utrecht. This means that we discontinue L+G-FR.26 in Live+Gov. [*Jij Maakt Utrecht*]	

Table 69: Live+Gov Functional Requirements and evaluation field trial results for 2nd project cycle

8 References

[1] P.A. Minnigh, et al., "Prototypes-demonstrators for first trials", Live+Gov Deliverable 5.2, November 2013.

[2] P.A. Minnigh, et al., "Detailed Use Case description", Live+Gov Deliverable 5.1, July 2013.

[3] F. Thiele, et al., "Technical verification and testing strategies", Live+Gov Deliverable 4.4, November 2013.

[4] H. Hartmann, C. Scheafer, and L. Niittylä, "Sensor Data Application", Live+Gov Deliverable 1.1, July 2013.

[5] L. Kovats, et al., "Conceptual documentation on issues, organization and stakeholder assessment", Live+Gov Deliverable 2.1, July 2013.

[6] F. Thiele, et al., "Integration concepts and guidelines", Live+Gov Deliverable 4.2, November 2013.

[7] Centraal Bureau voor de Statistiek, "Demografische kerncijfers per gemeente 2013", Centraal Bureau voor de Statistiek, 2013.

[8] D. K. Schneider and Vjollca Ahmeti, "Usability and user experience surveys", July 2013. Visited December 2013.
http://edutechwiki.unige.ch/en/Usability_and_user_experience_surveys

[9] Jacob Nielsen, "How Many Test Users in a Usability Study?", June 4, 2012. Visited December 2013. http://www.nngroup.com/articles/how-many-test-users

[10] Boris Blumberg, Donald R. Cooper and Pamela S. Schindler, "Business Research Methods", London: McGraw-Hill Education, 2008.

[11] F. Thiele, et al., "Report on Live+Gov toolkit requirements and architecture", Live+Gov Deliverable 4.1, July 2013.

[12] L. Kovats, et al., "Visualization of data injection from mobile sensing", Live+Gov Deliverable 2.3, January 2014.

[13] "Jekerkwartier als proeftuin voor nieuwe BuitenBeter-app", December 17, 2013. Visited January 2014. http://www.dichtbij.nl/maastricht/regionaal-nieuws/artikel/3274065

[14] Live+Gov Consortium, "Dissemination Plan", Live+Gov Deliverable D6.2, July 2013.

[15] J. Steensma, et al., "Webapplication for eGovernment Dialogue and Visualization - Public Beta, Functional Sketch", Live+Gov Working Paper, July 2013.

[16] A.F.M. Wierdsma, "Co-creatie van verandering", Delft: Eburon, 1999.

[17] C.K. Prahalad, V. Ramaswamy, "The future of competition: co-creating unique value with customers", Boston, Massachusetts: Harvard Business School Publishing, 2004.

[18] P. Henneman, "Burge(r)meesterboek: Methode voor lokale duurzame innovatie door radicale participatie", 2012. Visited February 2014.
http://screenbook.nl/burgermeesterboek

[19]   EUROCITIES, "Cities supporting eInclusion and citizen participation", November 2013. Downloadable via: http://www.eurocities.eu/eurocities/news/Cities-supporting-eInclusion-and-citizen-participation-WSPO-9DP7QU

[20]   "HTML 5 test – how well does your browser support html5?", Visited February 2014. http://html5test.com/results/desktop.html

[21]   L. Voogt, "Performance of JMU", Yucat research document, March 2013.