


| | |
|------------------|--|
| Project number: | 288577 |
| Project acronym: | urbanAPI |
| Project title: | Interactive Analysis, Simulation and Visualisation Tools for Urban Agile Policy |
| Instrument: | STREP |
| Call identifier: | FP7-ICT-2011-7 |
| Activity code: | ICT-2011.5.6 ICT Solutions for governance and policy modelling |

| | |
|------------------------|-------------------|
| Start date of Project: | 2011-09-01 |
| Duration: | 36 month |

| | |
|---|----------------------------------|
| Deliverable reference number and title (as in Annex 1): | D5.1: Assessment Methodology |
| Due date of deliverable (as in Annex 1): | 2012-08-31 |
| Actual submission date: | <i>see "History" Table below</i> |
| Revision: | |

| |
|--|
| Organisation name of lead contractor for this deliverable: |
| University of the West of England |

| | | | |
|---|--|---|---|
| Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013) | | |  |
| Dissemination Level | | | |
| PU | Public | X | |
| PP | Restricted to other programme participants (including the Commission Services) | | |
| RE | Restricted to a group specified by the consortium | | |
| CO | Confidential, only for members of the consortium (including the Commission Services) | | |

| |
|--------------------------------------|
| Title: |
| D5.1: Assessment Methodology |
| Author(s)/Organisation(s): |
| Zaheer Khan (UWE) |
| Working Group: |
| UWE, City partners, AEW, FH-IGD, AIT |
| References: |
| |

| |
|---|
| Short Description: |
| This deliverable sets out the evaluation methodology and planning to carry out evaluation of UrbanAPI tools and applications. |
| Keywords: |
| Testing, Evaluation, User Evaluation, Technical Evaluation, Requirements, Acceptance Criteria |

| History: | | | | |
|-----------------|--------------------|--------|---|------------|
| Version | Author(s) | Status | Comment | Date |
| 001 | Zaheer Khan (UWE) | new | Table of Contents, About this document, Introduction, Evaluation methodology, Evaluation stakeholders, Evaluation process, Operationalising the urbanAPI evaluation, Tools and Techniques | 15-06-12 |
| 002 | Zaheer Khan (UWE) | rfc | Revision, Summary and Conclusion and initiating consultation with urbanAPI WP5 partners | 19-06-12 |
| 003 | Zaheer Khan (UWE) | Final | Revised version based on consultation and feedback from Chiara Caranti (Bologna), Maria de Santiago (Vitoria-Gasteiz), Venco Bojilov (ASDE), Jens Dambruch (FH-IGD). | 31-07-12 |
| 004 | Zaheer Khan (UWE) | Final | Updated after quality review | 31-08-2012 |
| 005 | David Ludlow (UWE) | Final | Reviewed and prepared Public version | 12-09-2012 |

| Review: | | | |
|---------|---|--|------------|
| Version | Reviewer | Comment | Date |
| 001 | Jens Dambruch (FH-IGD), Cristina Garzillo (ICLEI) and Paola Mauri (AEW) | Document ready for quality review on 1 st August 2012. Quality review completed on 22 nd August 2012. Quality review process managed by Paola Mauri (AEW). | 22-08-2012 |

Disclaimer

This document is intended to be used for the FP7 urbanAPI project and is developed to meet specific needs of the project. It is accessible to everyone (i.e. non-project members) but the use of the concepts and contents of this document in other projects and/or context is solely responsibility of those project members or individuals. Authors and urbanAPI consortium will not be liable for the any negative results and damages or loss of resources due to use of the concepts and ideas.

Table of Contents

| | |
|---|----|
| About this Document | 5 |
| 1 Introduction | 5 |
| 2 Evaluation Methodology | 9 |
| 2.1 Background | 9 |
| 2.2 Evaluation perspectives | 9 |
| 2.3 The Methodology | 10 |
| 2.4 Evaluation Criteria | 12 |
| 3 The Evaluation process | 13 |
| 4 Stakeholders | 15 |
| 4.1 Evaluation Stakeholders | 16 |
| 4.2 Evaluation Stakeholders - Participation Planning Matrix | 19 |
| 5 Operationalising the urbanAPI Evaluation | 22 |
| 5.1 User Evaluation – Cycle 1 | 22 |
| 5.2 Technical Evaluation | 23 |
| 5.3 User Evaluation – Cycle 2 | 25 |
| 6 Summary of Evaluation Tools and Techniques | 27 |
| 7 Expected Outcome | 27 |
| References | 28 |
| Annex-1 – Evaluation Design Template for CIM2 | 30 |

About this Document

This document specifies the evaluation methodology as part of the *WP 5: Reality Check and Tools Validation* of the UrbanAPI project and is the Project Deliverable D5.1 – *Assessment Methodology*. The evaluation methodology consists of exercises to test, validate and verify the UrbanAPI tools and applications by involving various stakeholders. The main objective of this document is to indicate what will be evaluated? how will it be evaluated? when will the evaluation exercises be carried out? how will evaluation feedback to the development team be provided? who will participate in the evaluation exercises? and what are the expected outcomes of the evaluation? The intended audiences for this document are UrbanAPI project partners, local city stakeholders, UrbanAPI stakeholder board and other stakeholders who are not directly involved in the project but can benefit from the developments and outcomes of the project. These external stakeholders can be urban planners, policy and decision makers, general public, ICT developers, environmentalists, NGOs and related projects in the domain of urban planning, governance and policy making. A more detailed list of stakeholders can be found in the chapter 4 of this document.

1 Introduction

This document specifies the overall evaluation methodology and the process to evaluate the urbanAPI project as part of the work package 5 (WP5) activities. In this document we elaborate on the overall planning of the evaluation, the evaluation methodology of the project, the evaluation process, main actors to carry out the evaluation, evaluation tools and techniques and the expected outcomes.

UrbanAPI is an ICT driven collaborative research project, built upon the foundations of the smart cities agenda, which aims to support activities such as issue identification, policy analysis, consultation, decision and evaluation in urban planning and land management policy. It focuses mainly at the urban core (i.e. neighbourhood level), city level and city-regional scale by developing three applications consisting of 3D Virtual Reality, Public Motion Explorer and Urban Growth Simulation. Four European cities (Vienna, Bologna, Vitoria-Gasteiz and Sofia) are participating in the project as pilot case studies for the implementation of the above applications. A detailed set of user requirements have already

been acquired from the above cities as part of the work package 2 activities and have been documented in Deliverable *D2.1 – User requirements definition*. These requirements provide basic scenarios, user needs and requirements to develop essential criteria for preparing the evaluation exercises for the UrbanAPI applications. Each city is participating in at least two different applications to identify and cross-verify commonalities and generic aspects. The following table 1 depicts the city participation in different urbanAPI applications based on the local policy priorities.

Table 1: urbanAPI applications for cities

| | 3D VR | Public Motion Explorer | Urban Growth Simulation |
|-----------------|-------|------------------------|-------------------------|
| Vienna | √ | √ | |
| Bologna | √ | √ | |
| Vitoria-Gasteiz | √ | √ | |
| Sofia | | √ | √ |

Before we elaborate the evaluation methodology for the UrbanAPI project in detail, we would like to discuss what is meant by evaluation in the UrbanAPI project. Wikipedia summarizes various approaches to conduct the evaluation and defines evaluation as:

“Evaluation is the comparison of actual impacts against strategic plans. It looks at original objectives, at what was accomplished and how it was accomplished. It can be formative, that is taking place during the life of a project or organization, with the intention of improving the strategy or way of functioning of the project or organisation. It can also be summative, drawing lessons from a completed project or an organisation that is no longer functioning. ... ” [1].

In the above context, figure 1 below depicts a holistic view of the UrbanAPI evaluation. It shows that there are different stakeholders involved in the development and evaluation of the project including city administrations, general public and technology experts, etc. These stakeholders participate in defining the specific requirements for the development of a generic UrbanAPI Toolset. Further, acceptance criteria are also defined based on which specific capabilities of the UrbanAPI Toolset and applications will be assessed. The UrbanAPI toolset will be used for the development of three

applications (3D VR, GSM and Interactive Urban Simulation) at different urban governance scales. These tools and applications will be evaluated using both technical and user evaluation techniques and feedback will be provided to stakeholders to repeat the development process.

As the UrbanAPI project has adopted an agile development methodology (i.e. development process running with iterative and concurrent activities), the main emphasis here will be to ensure formative evaluation will be carried out at different development stages of the project to provide timely feedback to the development team. Also a summative evaluation will be conducted at the end of the project mainly to elaborate project development experiences, lessons learnt, pros and cons and suggested improvements.

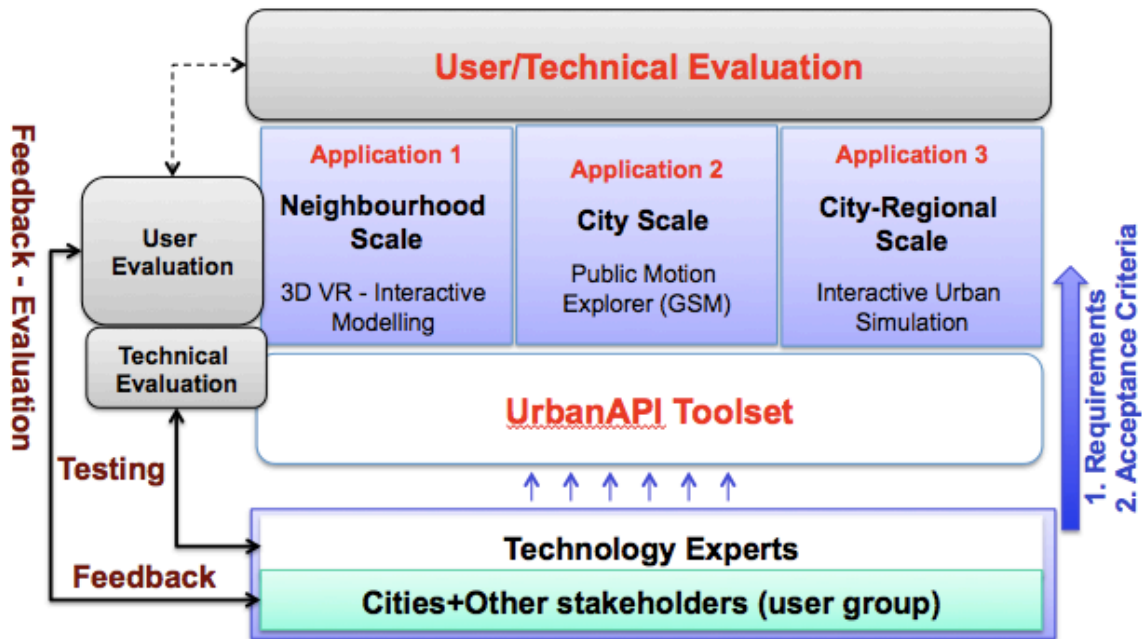


Figure 1: Overall Evaluation in the UrbanAPI project

The overall project evaluation is divided into two major categories: i) User evaluation; and ii) Technical evaluation, as depicted in Figure 2.

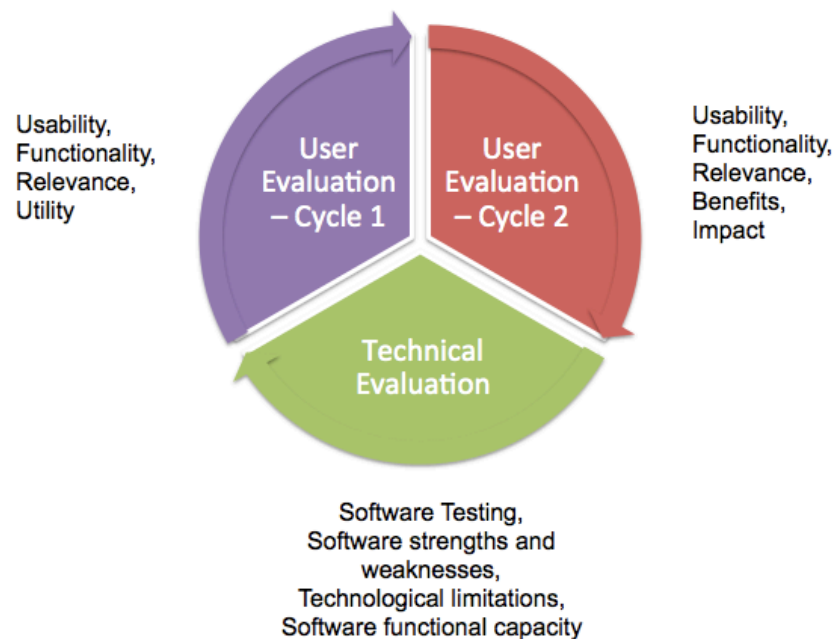


Figure 2: UrbanAPI Evaluation categories

The technical evaluation aims to test UrbanAPI tools and applications and cross verify the outcomes with stated requirements specifications. Mainly it focuses on the software quality aspects and intends to identify technological strengths and limitations to fulfil stakeholder requirements. Technical evaluation mainly aims to use classical testing approaches to find errors and bugs in the software solutions and to ensure that functional and non-functional requirements are satisfied. Technical evaluation will be performed for both development cycles of the project and the results will be documented in deliverable *D5.3: Technical Evaluation*. On the other hand, user evaluation is performed to assess the relevance, usability, functionality, benefits and overall impact on participatory policy modelling by using the ICT solutions developed in UrbanAPI. User evaluation will be carried out in two phases or cycles as per the agile methodology of the project and the results will be documented in the deliverable reports: *D5.2: User Evaluation – Cycle 1*, and *D5.4: User Evaluation – Cycle 2*.

City partners, local city stakeholders and project partners will actively participate in the successful application of the evaluation exercises in UrbanAPI. This document serves as a roadmap to conduct various evaluation activities during the lifetime of UrbanAPI and results will be used to complete various work package 5 deliverables.

2 Evaluation Methodology

2.1 Background

One of the challenges in the development of a collaborative and integrated project (hereafter referred as UrbanAPI) is to collect and analyse the requirements from different stakeholders in order to develop the required software and applications driven by more coherent and validated requirement specifications. Furthermore, the evaluation of outcomes against the stated objectives and requirements of a project requires a structured and coherent evaluation methodology. Undertaking the evaluation of a collaborative project is not straightforward due to a variety of factors, for instance, the timely delivery of different components of the system when: i) there are dependencies between various components; ii) several stakeholders from different countries are involved who typically have diverse technological and/or application domain backgrounds and objectives; iii) there exist multiple and sometimes conflicting application specific and user-defined requirements; and iv) there are multiple and often conflicting technological research and development objectives. All these aspects require a coherent evaluation methodology and planning to adapt to the diverse requirements of collaborative integrated research projects.

In the above context, notwithstanding the fact that there are many research and system evaluation approaches, methods and techniques [2-9], the assessment of the outcomes of an integrated and collaborative project using ad-hoc evaluation approaches may not fully reveal the limitations and shortcomings of the underlying system. Also, a single approach or technique may not be fully appropriate in assessing the various evaluation aspects of the system. Therefore, it is crucial to define and/or select the best-fit evaluation methodology and techniques for the evaluation of such systems. In this regard in the following chapters we introduce the overall evaluation methodology of UrbanAPI.

2.2 Evaluation perspectives

According to Juristo N. et. al. [7], 'the evaluation of a system can be divided into four basic types:

- Checking the correctness of the system structure for design and coding errors (i.e. *verification*).
- Checking the correctness of the content or semantics of the system (i.e. *validation*).
- Checking the system externality for its operational success (i.e. *system usability*).

- Checking what improvements the current system has brought to the organisation (i.e. *usefulness*)’.

Based on the above types and nature of the development of the UrbanAPI toolset and applications which are both user-driven (based on actual user requirements) and technology-driven (based on a state of the art analysis), different perspectives are addressed in the evaluation of the UrbanAPI outcomes. Thus the evaluation results are based on the two perspectives: a more technology-oriented i.e. “perspective A: software quality” and a more user and context-oriented “perspective B: relevance and benefits”. This may also be referred as technical evaluation and user evaluation, respectively as discussed in chapter 1.

Perspective A looks at software quality including usefulness, functionality, reliability and usability. This perspective should answer the prime question: “Does the newly developed software meet the functional and non-functional requirements defined in the specification?” Perspective A is based on the main evaluation criteria as defined in the ISO 9126 quality model including maintainability, portability, functionality, reliability, usability, efficiency etc.

Perspective B assesses the outcomes of the UrbanAPI project from the perspective of benefits to potential users and relevance to the stated or implied objectives of the UrbanAPI project. This perspective is particularly important for the city scenarios for specific UrbanAPI applications and should answer the prime question: “To what extent do the UrbanAPI ICT toolset and applications achieve the stated and implied objectives of the project such as participatory policy development, public participation for planning initiatives, facilitation for planning decisions, etc?” Relevance, benefits and overall impact are the main evaluation criteria under Perspective B.

2.3 The Methodology

In order to perform the above evaluation in UrbanAPI, the Criteria Indices Metrics – Ver 2 (CIM2) methodology has been introduced which primarily defines the means to secure evaluation results as shown in Figure 3. CIM2 is extended from the experience developed in the FP6 HUMBOLDT project [10] [16]. Basically, this approach considers the design of the evaluation to be carried out, and defines a set of criteria based on specific aspects (e.g. user functional and non-functional requirements), which

need to be considered, for instance aspects related to usability, functionality, performance, etc. Each criterion may have additional associated sub-criteria to address various other aspects. Each sub-criterion is made operational by defining one or more indicators which address the evaluation criteria in question. In order to better understand the context of the evaluation for a particular aspect, each indicator is represented by one or more questions. Also, for each indicator (and associated questions) some metrics are defined to judge whether or not the result is regarded as good or bad. Additionally, qualitative assessment is also included in CIM2 to enable evaluators to provide subjective (and/or objective) assessment mainly concerning the benefits, relevance and the overall impact of the UrbanAPI ICT solutions. These qualitative outcomes can be in the form of subjective statements. For instance, *‘the 3DVR application improved accessibility to urban planning activities and significantly improved public participation for collaborative decision making.’* Techniques like system usability scale [17] or its variations [18] may be utilised to give a global view of the subjective assessment.

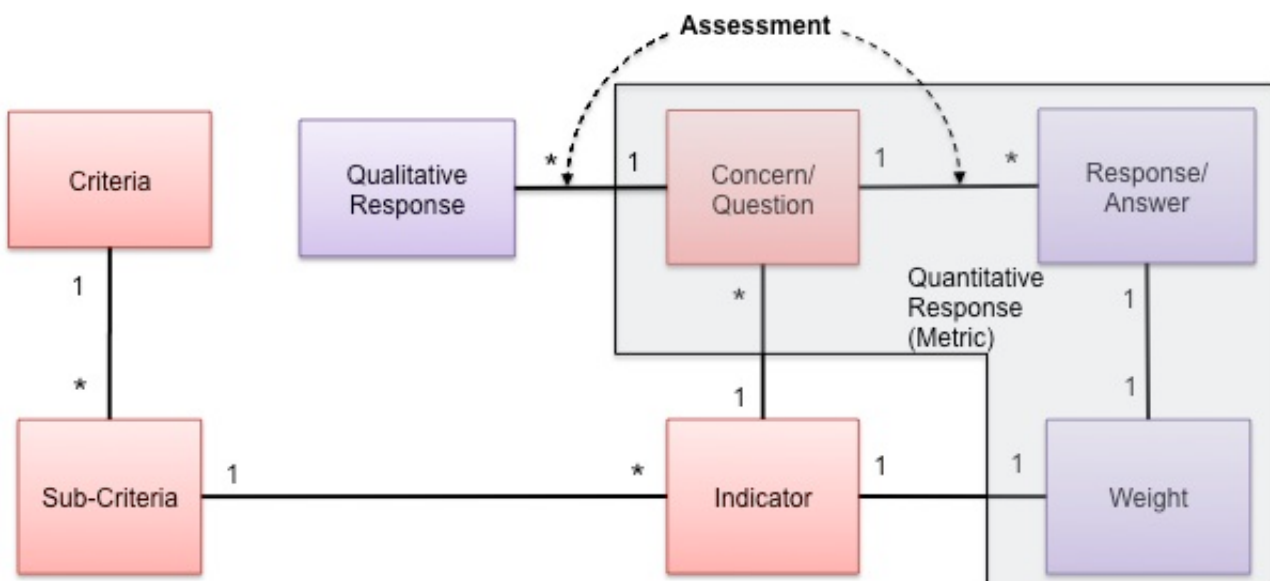


Figure 3: Evaluation Methodology: Criteria-Indicator-Metrics - 2

As an example, we illustrate in Figure 4 ‘usability’ as one of the main elements derived from ISO 9126 (ISO/IEC 9126-1 2001), and selected as a main criterion with two further sub-criteria defined as: ‘learnability’ and ‘understandability’. As an example, each sub-criterion identifies indicators, also considered as concerns to be evaluated. In order to define metrics for indicators specific questions are defined which address the aspects to be evaluated for each indicator. These questions can have specific response options each defined with quantitative and measurable weights. *Annex-1* provides

an evaluation design template derived from CIM2 to be used for indicating overall criteria, indicators, metrics and available resources.

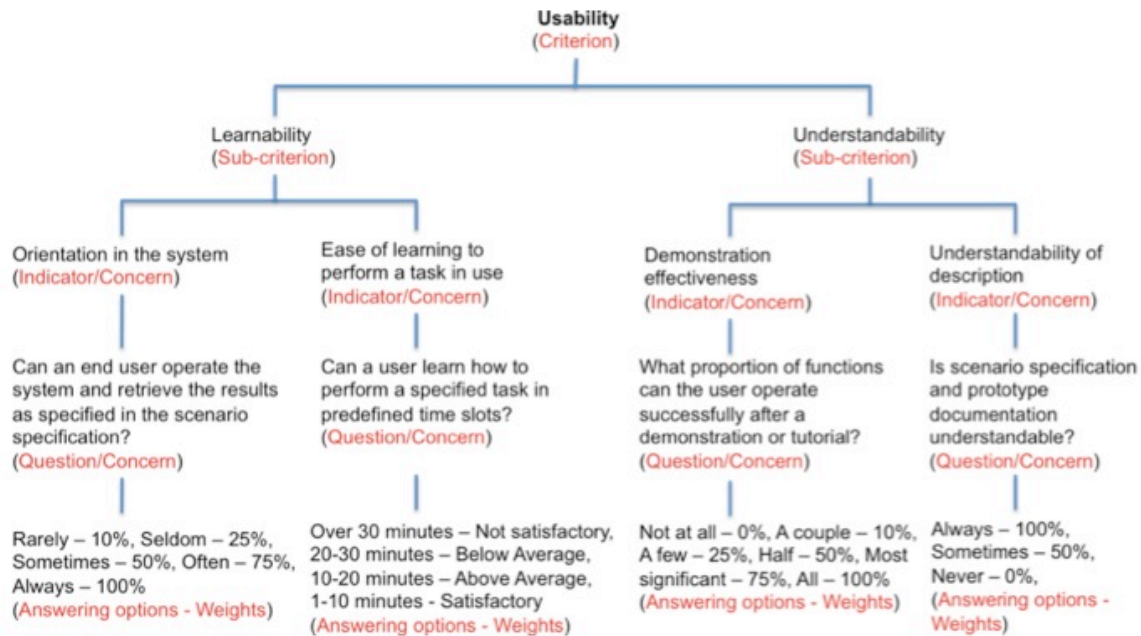


Figure 4: CIM2 methodology: An Example

2.4 Evaluation Criteria

In general, CIM2 identified above adopts both white-box (Technical evaluation) and black-box (User evaluation) testing approaches. The main purpose of the evaluation is based on the following criteria from both A and B evaluation perspectives:

- **Relevance** – The relevance of the UrbanAPI tools and applications in achieving the *project goals* for the stakeholders and for illustrating the goals and potentials of ICT enabled *participatory urban planning and governance, policy making process, information flow between different levels of governance*, etc;
- **Utility and Benefits** – The potential benefits of the application scenarios in *local planning process, commonalities, inspirations, comparisons, public engagement*, etc;

The software quality perspective was also considered using the following criteria:

- **Usability** - The capability of the UrbanAPI tools and applications to be *understood, learned, used and appreciated by the stakeholders/end-user*, when used under specified conditions;

- **Functionality** - The capability of the UrbanAPI tools and applications to provide *functions which meet stated and implied needs, or requirements* when the software is used under specified conditions.

And finally,

- **Impact** - The overall effect in *public engagement in policy decision making processes, public behavioural change, awareness, intervention, improvement and acceptance* of planning initiatives, etc.

3 The Evaluation process

Figure 5 depicts the overall process to enact the project evaluation. Here, circles depict different stages and arrows show transition from one stage to another. A self-directed arrow means that this stage may possess iterative behaviour. These stages are explained as below:

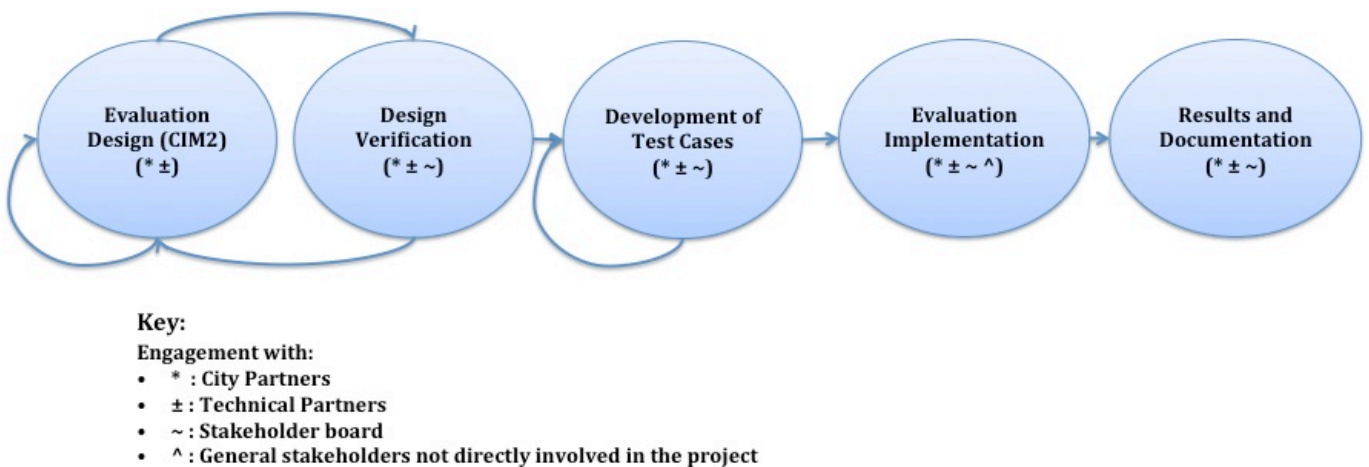


Figure 5: The evaluation process

1. Evaluation Design: this stage is used to prepare the evaluation contents (i.e. criteria, indicators, questionnaire, etc) as per the template depicted in *Annex-1* and consists of the following activities:

- a) Identification of the main set of criteria based on the requirements specification, application objectives and stakeholder needs.

b) Deriving sub-criteria and identifying indicators to achieve the objectives of the project and its applications.

c) Deriving questionnaire and response options with associated weights to measure the outcomes.

d) Identifying the means to perform specific evaluation exercises to answer the evaluation questions.

2. Design Verification: this stage is used to verify the evaluation design by engaging with stakeholders and consists of the following activities:

a) To verify that the evaluation design contents (i.e. criteria, questionnaire, etc) capture important assessment elements and are in line with stakeholder needs and requirements.

b) Highlighting the importance (high, medium, low) of specific questions.

c) Indicating improvements in the questionnaire, response options, metrics etc.

d) Identification of additional criteria, indicators, questionnaire and metrics.

3. Development of Test Cases: this stage is used to develop test scenarios to validate software in meeting user requirements and needs. The test cases will be derived from requirements definition with the objective to apply specific CIM2 questionnaire.

4. Evaluation Implementation: in this stage the actual evaluation is carried out by using different mechanisms such as web-based tools, workshops, etc. Evaluation stakeholders will be actively involved in performing the evaluation.

5. Evaluation Results and Documentation: all the above stages including evaluation results are properly documented, compared with related projects (i.e. benchmarking) and verified by the stakeholders.

4 Stakeholders

Bryson J, et. al. [14] emphasised proper identification of different categories of evaluation stakeholders to be utilised for evaluation planning, design, implementation and decision making processes. UrbanAPI stakeholders refer to an individual or an organisation or any legal entity that have vested interest (i.e. financial, technical, knowledge, policy etc) in the outcomes and/or success of the project. Stakeholders can be individuals who are involved in the development of the system (i.e. operators, developers or maintenance staff) or use (i.e. functional or operational beneficiaries) the system or who are mainly interested in the outcomes of the project (political or financial beneficiaries). In UrbanAPI , stakeholders play a key role in deriving requirements for the development of tools and applications, data provision and validating the outcomes of the project. Figure 6 depicts a high level stakeholders diagram from the UrbanAPI perspective:

- Local city stakeholders are at the centre of the project development process. This includes city municipality/administration and various planning and IT departments, policy makers, citizens, etc.
- Local city stakeholders are supported by technology and domain experts i.e. IT experts, 3D application experts, GIS experts, urban planners, transport planners, architects, developers etc.
- Stakeholder board consists of selected European user related organisations and city networks that have research objectives similar to the UrbanAPI project and are interested in the outcomes of the project development and applications. The Stakeholder board not only aims to link UrbanAPI development from local to regional and European scales and provide expert opinion, but also to evaluate the outcomes of the project applications.
- General public are the indirect beneficiaries of the project outcomes.



Figure 6: UrbanAPI Stakeholders: Simplified Concept

4.1 Evaluation Stakeholders

The above stakeholder diagram (Figure 6) is further populated with more concrete stakeholder roles by using the ‘Onion Model’ approach, indicating political beneficiaries, functional beneficiaries, normal operators and support operators [11]. Using the structure of the product onion model [11], Figure 7 depicts the detailed relationship of stakeholders for UrbanAPI. Each circle in the onion diagram denotes a subset of entities relevant to the development of the project, for example, specific roles (as ‘actors’ in UML Use cases) such as Policy Experts, Domain Experts and IT Experts in ‘The System’ circle.

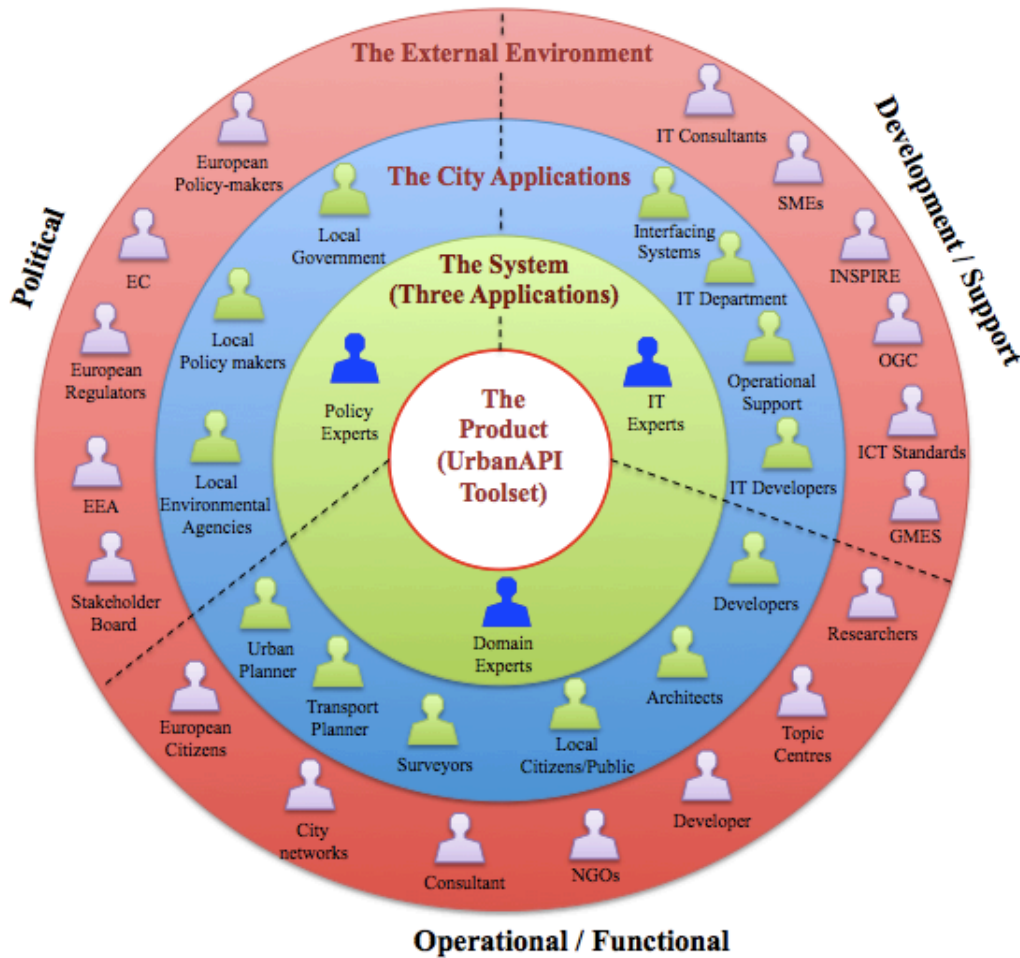


Figure 7: UrbanAPI Stakeholders based on the Onion Model

The UrbanAPI stakeholder onion diagram consists of four concentric circles:

- a) *'The Product'* is the inner circle referring to the UrbanAPI Toolset that provides basic functionality for the development of the UrbanAPI applications.
- b) *'The System'* refers to the three applications of UrbanAPI that includes the UrbanAPI Toolset and its human operators, local policy or domain specific regulations and standard operating procedures, etc.
- c) *'The City Applications'* refer to 'The System' plus any human beneficiaries (e.g. primary and secondary stakeholders).
- d) *'The External Environment'* refers to 'The City Applications' and any other stakeholders (e.g. Tertiary stakeholders).

The onion diagram is useful in indicating different stakeholder roles, and can be used to apply 'Stakeholder role play' techniques [15, p.133] to protect the interests of stakeholders who are under-represented or difficult to access [14].

In order to have a more general stakeholder classification, in UrbanAPI the following three categories are adopted from [12]. This classification is performed to maintain consistent use of stakeholder roles in different work packages (e.g. WP2 – requirements and WP5 – evaluation) of the project. All stakeholders can be financial beneficiaries and the proposed stakeholder categories are:

1. Primary stakeholders

These are end users who will actually operate the software application to visualise specific scenarios e.g. urban planners, transport planners, architects, surveyors, IT operators, developers and maintenance operators etc. They can also be referred as *primary key evaluation stakeholders* who can participate in the evaluation and can influence the evaluation outcomes by representing different dimensions of stakeholders interests, priorities, needs, concerns, powers and perspectives [14].

2. Secondary stakeholders

This category refers to the functional beneficiaries e.g. those who consume system outputs in their work but do not actually operate the system e.g. architects, transporters, real estate, construction/developers, SMEs/executives, transport/urban planners, general public, etc. They can also be referred to as potential *candidate or intended evaluation stakeholders* [14] who have interest in the evaluation outcomes and can provide useful inputs in planning and designing the evaluation but may not have capacity to participate in the evaluation.

3. Tertiary stakeholders

This category refers to functional or political beneficiaries of the information generated by the system supporting strategic planning and decision-making e.g. local government, policy-makers, regulators, general public, etc. Such stakeholders are considered to have power (political influence, decision-making etc) and interest (advancement in policy making processes using ICT, participatory governance, etc) that can significantly influence the evaluation process in general and have impact on the technological adoption and future policy implementation in particular. They can also be referred as

intended evaluation stakeholders [14] who are interested in the evaluation outcomes but may not have capacity to participate in the evaluation.

4.2 Evaluation Stakeholders - Participation Planning Matrix

Bryson J, et. al. [14] provide a composite of techniques to engage stakeholders in different stages of the evaluation process. The participation planning matrix technique (originally from International Association for Public Participation) appears to be useful in indicating the level of participation from various stakeholders at different stages of the evaluation process. We aim to adapt it to indicate the level of engagement by different stakeholders at different stages of the evaluation process in the project as presented in table 2 below. The table indicates different levels of engagement in project development and evaluation. In table 2, the *type of involvement* indicates both benefits to the project as well as privileges to different types of stakeholders based on their level of engagement. Furthermore, *the overall evaluation process* (in table 2) identifies the level of engagement by different types of stakeholders at different stages of the evaluation process (i.e. evaluation design – planning and CIM2, test scenarios development, evaluation implementation – user and technical evaluation, results and documentation).

Table 2: Participation Planning Matrix adopted from [14]

| Levels of Engagement | | | | | | | |
|---|----------------|--|--|---|--|---|---|
| Do not Engage | As Data Source | Inform | Consult | Involve | Collaborate | Empower | |
| Types of Involvement | | | | | | | |
| <i>Commitment from Project and privileges</i> | No promises | Fulfil data usage protocols; quality assessment; identification of missing data | Inform project outcomes and evaluation results | Consider input for evaluation design and inform evaluation output | Include interests and inputs in evaluation design; considered part of the evaluation decision-making process | Include advice and suggestions to greatest extent possible; considered part of the evaluation decision-making process | Fully engaged in the evaluation planning, design, implementation and decision-making process |
| <i>Benefits of engagement</i> | Not involved | Provides necessary data; assessment of fulfilment of application specific data needs | Dissemination and wider user community | To get different perspectives, issues, needs, concerns | Attracting attention and establishing credibility of evaluation results | Act as primary intended users and have great interest and availability and can influence evaluation outcomes | Capacity development; to be able to participate in evaluation and have sense of ownership of the evaluation |
| The Overall Evaluation Process (Stakeholder Categories) | | | | | | | |
| | | Primary, Secondary & Tertiary | Secondary & Tertiary | Secondary & Tertiary | Secondary & Tertiary | Primary & Secondary | Primary & Secondary |
| <i>Evaluation Design (i.e.</i> | | | | | | | IT developers, city partners |

| | | | | | | | |
|--|--|--|--|--|---|--|---|
| Planning evaluation exercises) | | | | | | | including urban and transport planners |
| <i>Design Verification</i> (i.e. CIM2 based contents) | | | | | Stakeholder board | City partners including urban and transport planners and IT department | Urban and transport planners, architects, IT developers, maintenance operators, |
| <i>Development of Test Cases</i> | | Architect, urban and transport planners, domain experts, IT developers | | Local SMEs, Domain experts | Stakeholder board, local city stakeholders including environmentalists, NGOs, etc | IT developers, maintenance operators, urban and transport planners, IT department, stakeholder board | Urban and transport planners, architects, developers, IT developers, maintenance operators, environmentalists etc |
| <i>Evaluation Implementation (i.e. Technical Evaluation)</i> | | | Domain experts | | | IT developers, Local city IT department | IT developers and city partners mainly IT department |
| <i>Evaluation Implementation (i.e. User Evaluation)</i> | | General public | EC, Local city stakeholders, e.g. general public | Local city stakeholders e.g. NGOs, environmentalists, general public | Domain experts, local decision makers, policy makers, and general public | Stakeholder board, local agencies not directly involved in the project e.g. transport department, | City partners including urban and transport planners, architects, IT department, |
| <i>Results and Documentati</i> | | | Local city partners, | | Stakeholder board and local | Stakeholder board | IT developers, City partners |

| | | | | | | | |
|--|--|--|--------------------|--|---|--|--|
| on (i.e. Reporting D5.1, D5.2, D5.3, D5.4) | | | general public, EC | | stakeholders not directly involved in the project | | including urban and transport planners, architects, IT department. |
|--|--|--|--------------------|--|---|--|--|

5 Operationalising the urbanAPI Evaluation

5.1 User Evaluation – Cycle 1

User evaluation will be performed to validate the end-to-end outcomes of the UrbanAPI applications against the stated and implied objectives and application scenario requirements of the project. The overall aim is to validate functionality, usability and relevance of the applications and identify benefits and overall impact of the application. One common requirement in the urbanAPI applications is to ensure accessibility to applications via web browsers. This feature enables various local stakeholders to participate in the application evaluation exercises. As the main objective here is to perform end-to-end (i.e. black-box) evaluation, the requirements specifications including application scenarios, use-cases and requirement definitions provide necessary information to build test scenarios to satisfy stakeholder needs. A web-based interface to record the evaluation results enables wider participation from local stakeholders and automated analysis of the evaluation outcomes.

Bryson J, et. al. [14] listed techniques from literature to analyse, interpret and to prepare recommendations based on the evaluation outputs. These techniques can be '*evaluation recommendation support versus opposition grids*' and '*recommendation attractiveness versus stakeholder capability grid*'. The former technique helps in identifying the level of importance of the recommendations to the stakeholder against the support, opposition or neutrality. The later technique helps in assessing which recommendations are likely to be implemented or disregarded due to capacity, resources and technical issues [14]. The user evaluation cycle 1 will adopt the above techniques to provide feedback to be utilised for the implementation of UrbanAPI software and applications for development cycle 2.

Figure 8 depicts overall planning for the user evaluation cycle 1. It shows that evaluation for each application will have three main activities: i) *Preparation* – i.e. designing evaluation with specific test scenarios and questionnaire; ii) *User evaluation* – i.e. implementing or carrying out the actual evaluation; and iii) *Reporting* – i.e. documenting the evaluation outcomes that will contribute to Deliverable D5.2 – *User Evaluation – Cycle 1*. Furthermore, outputs from each evaluation exercise will be used: i) as feedback to the development team indicating strengths and limitations of software applications as well as new emerging requirements; and ii) as an input for the preparation of user evaluation – cycle 2.

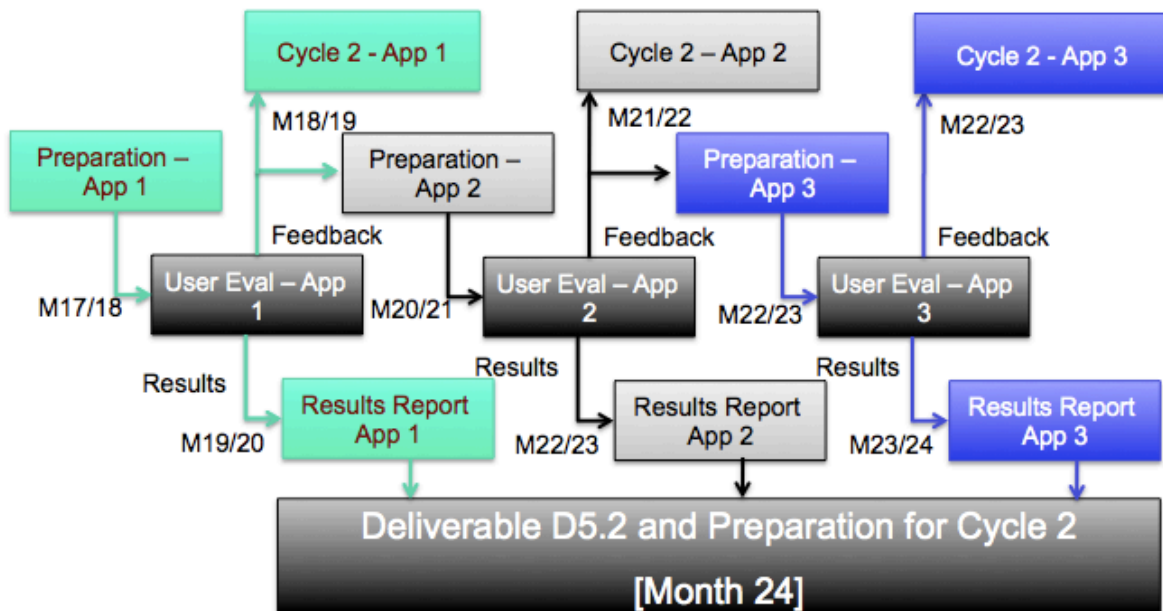


Figure 8: Planning for User Evaluation – Cycle 1

5.2 Technical Evaluation

Technical evaluation mainly refers to software testing against stated and implied functional and non-functional requirements of the UrbanAPI application scenarios. It includes both verification and validation aspects. In general the verification element ensures that the software application has been built according to the requirements and design specifications. Furthermore, validation ensures that the software application actually meets the user's needs, and that the requirements and respective design specifications were correct in the first place. The verification element responds to the question 'Are we building the product right?' i.e. requirements are correctly mapped to design and

implementation components. It is like phase testing of the software development process with stated testing conditions for each phase. Verification results are complemented with the validation element. The validation here responds to the question 'Are we building the right product?' i.e. the output of the software testing conforms to the expected behaviour as detailed in the acceptance criteria for application requirements. Here validation will be performed at the end of the development process. The main objectives are to test functional as well as quality attributes i.e. to identify software faults (i.e. incorrect functionality), bugs (i.e. errors), malfunction (i.e. missing or wrong functionality), etc.

In addition, the technical evaluation will assist in verifying the following ICT and policy modelling objectives of the project:

- proof the conceptual integrity of the meta-model and the rule languages hierarchy;
- show the power of the rule-based approach by capturing the data integration model, data requirements, constraints and presentation rules for the applications in each city;
- conduct load testing, unit testing and integration testing, continuous integration with code coverage, performance analysis and other technical measures to show the technical correctness and performance of the toolkit and the applications;

To have a more structured testing, a 'unit testing' approach will be adopted and developers and testers may consider different tools [13] and techniques to derive test cases from the logic of the program: boundary value analysis, error conditioning, equivalence partitioning, etc [19], continuous integration for automated testing [20], code analysis tools [21] for code coverage [22], etc. Similarly, feedback on the individually developed components will be acquired from the stakeholders to validate the outcomes. As the requirements specifications for the development of the urbanAPI tools and applications are structured to explicitly identify the acceptance criteria against each requirement statement, this would be used to achieve the above objectives of technical evaluation.

Based on the urbanAPI agile development methodology, the technical evaluation will be performed in two stages (or cycles). Figure 9 depicts the overall lifespan of these evaluation stages contributing to the Deliverable *D5.3 – Technical evaluation*. In Redmine (for requirements management – D2.1: Requirements Definition) requirements for each application will be mapped to different issues which will provide traceability to design and implementation as well as a basis for unit testing and validation. Based on the outcomes of the cycle 1 evaluation (i.e. faults, bugs, malfunctioning, emerging

requirements etc), cycle 2 evaluation aims to cover a comprehensive technical evaluation before user evaluation cycle 2. The main outcome of this evaluation will be in the form of a report including i) test cases and testing performed by developers/testers; ii) results of the testing; iii) tests or feedback or identification of issues with software by the UrbanAPI user community; iv) steps taken to resolve issues identified in iii; v) technical strengths of the UrbanAPI software; vi) technological limitations to accomplish user requirements; and vi) suggestions and recommendations to mitigate technological limitations. All these outcomes will be documented and contribute to *Deliverable D5.3 – Technical evaluation*.

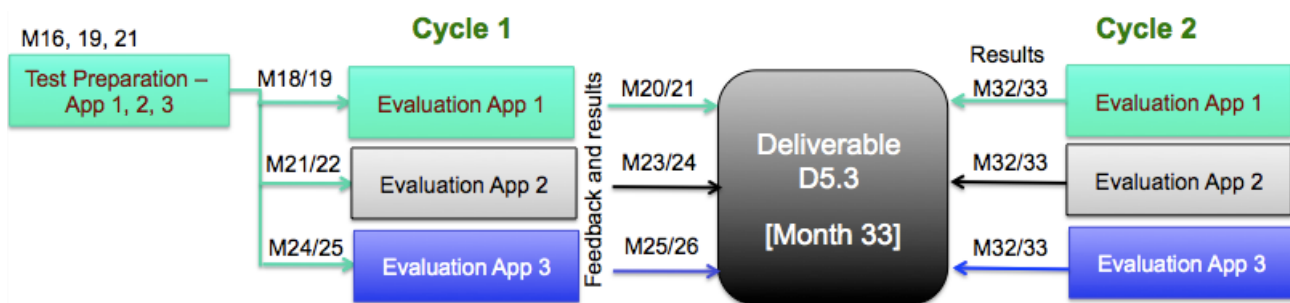


Figure 9: Planning for Technical Evaluation

5.3 User Evaluation – Cycle 2

User evaluation – cycle 2 will be performed on the final products developed by the project. Cycle 2 will extend the evaluation exercises based on the experiences, recommendations and results from the user evaluation cycle 1. In general, cycle 2 will complement the results obtained in the user evaluation cycle 1 as reported in *Deliverable D5.2- User evaluation – Cycle 1*. In this cycle, evaluation against all the criteria (usability, functionality, benefits, relevance and impact) will be performed. In addition to the techniques indicated in the user evaluation – cycle 1, the cycle 2 evaluations can perform a state-of-the-art literature review of the evaluation of collaborative projects with similar thematic domains (i.e. use of ICT for citizen participation, governance, urban planning and policy development). The analysis of the literature will provide a foundation for the benchmarking and comparative analysis of the evaluation results from the project that will yield benefits, relevance and assess the overall impact of the UrbanAPI tools and applications. Furthermore, Bryson J, et. al. [14] presented ‘*recommendation implementation strategy development grid*’ technique from the literature that can be adapted to

indicate strengths, limitations and required resources to accomplish action plans based on stakeholders interests and available resources.

Figure 10 shows how the evaluation workshops will be carried out in each partner city in order to gain wider participation from local stakeholders. In addition, web-based evaluation will also be used to obtain quick feedback. All these inputs will contribute to Deliverable D5.4 – User Evaluation – Cycle 2.

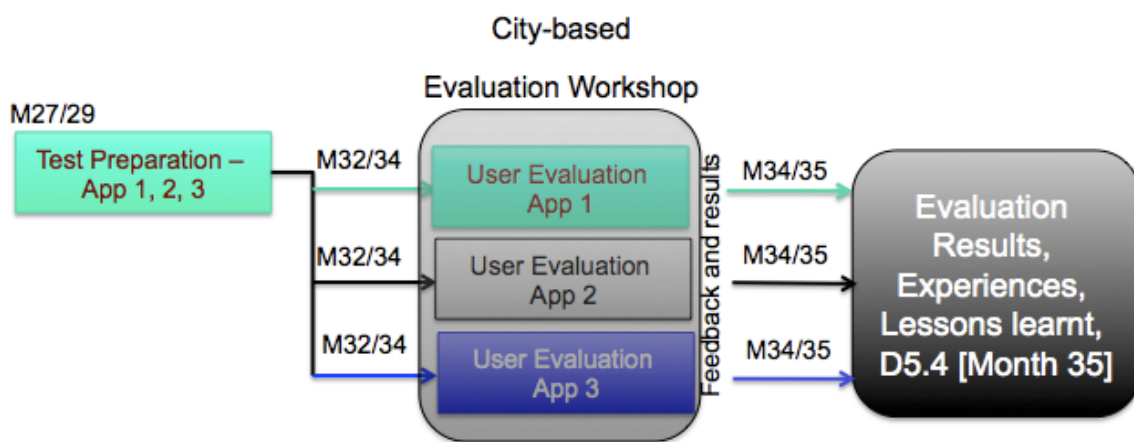


Figure 10: Planning for User Evaluation – Cycle 2

In general, the above user evaluation (cycle 1 and cycle 2) will enable the evaluation stakeholders to verify the following policy making and public participation objectives:

- indirect (through usage data, such as number and type of interactions – i.e. application specific test cases) and direct (through CIM2 questionnaire) evaluation of qualitative and quantitative aspects of the implemented system by all city stakeholders, for example trust in the solutions;
- analysis of application specific and city based policy making processes on the basis of the agile process model to collect and share lessons learned.

In addition, the following ICT and policy modelling objectives will also be verified:

- conduct usability studies to analyse the impact of the planned 3D User Interface elements, the 3D cartographic approaches and the visual analytics approaches – Fraunhofer has a complete Usability lab (including the multi-touch table) that can be used for this purpose;
- conduct task-based usability studies to show the effectiveness/efficiency of the rulebased approach and the corresponding user interfaces.

6 Summary of Evaluation Tools and Techniques

This chapter lists various tools and techniques which can be used to facilitate evaluation exercises. For example:

- **Web-based accessibility:** for user evaluation, a web-based interface will be used to facilitate participation for a wider audience.
- **Evaluation workshops:** specific evaluation workshops can be organized in participating cities.
- **Unit Testing:** for technical evaluation, unit testing (e.g. coverage analysis) will be performed using classical software testing techniques like boundary value analysis, error conditioning, etc to assess the quality and functionality of the software.
- **Evaluation guideline:** a guideline with detailed instructions how to perform evaluations will be provided to evaluation stakeholders.
- **Scenarios/Test case:** will be derived from requirement scenarios, use cases and requirements definitions.
- **Accessibility to UrbanAPI tools and applications:** a guideline with detailed instructions from where to access the urbanAPI tools and applications and test data.
- **Evaluation questionnaire:** evaluation questionnaire developed by using CIM2 methodology.
- **Quantitative (statistics, graphs etc) and Qualitative assessment:** Both qualitative and quantitative methods will be used to critically assess the evaluation outcomes.
- **Tools and Techniques:** different techniques presented in [14] can be adopted. More specifically, *role-play* techniques to ensure that the interests of under-represented or difficult to access stakeholders are captured and represented; *purpose network* technique to identify common and conflicting evaluation purposes that may be considered as concerns or criteria to be used in CIM2. Also, *evaluation recommendations support versus opposition grids* and *recommendation attractiveness versus stakeholder capability grid* may be considered for evaluation cycle-1 analysis purposes. Similarly, *recommendation implementation strategy development grid* can be used to depict overall assessment of limitations, strengths and resources needed to accomplish stakeholders' interests and goals based action agenda.

7 Expected Outcome

In this document, the CIM2 is presented in detail that helps in identifying the criteria, indicators and metrics to assess whether or not the stated and implied objectives of the project have been achieved.

Furthermore, the overall evaluation process to operationalise the evaluation of the urbanAPI tools and applications is presented. As a result of applying the above CIM2 methodology the expected outcomes will be related to, but not limited to, the following aspects:

- **Software Quality:** ability of the software to be reliable, secure, maintainable, etc.
- **Functionality:** ability to respond to the stakeholder requirements.
- **Relevance:** ability of the software to respond to needs of the stakeholders.
- **Utility and Benefits:** ability of the software to add value to organizational processes and citizen engagement for planning initiatives, policy-making and collaborative decision-making.
- **Usability:** ability of the software to be understood, executed and used by various stakeholders.
- **Limitations:** indicating the technological limitations or conflicting stakeholder needs.
- **Experiences and Lessons Learnt:** reflections on the overall development process and what could have been done better to achieve additional results.
- **State of the art gap analysis:** compare and contrast the UrbanAPI solutions with other existing products.
- **Impact:** overall effect of the software to different stakeholders, planning and policy-making processes, socio-technical innovation, etc.

The above outcomes will be documented as an input to the following deliverables:

- D5.2: User Evaluation Cycle 1
- D5.3: Technical Evaluation
- D5.4: User Evaluation Cycle 2

References

- [1] Evaluation: Wikipedia definition, URL Access: <http://en.wikipedia.org/wiki/Evaluation>
- [2] Andras P. (2011), 'Research: metrics, quality, and management implications', *Research Evaluation*, 20/2: 90–106.
- [3] Babar MA, Zhu L, Jeffery R. (2004), 'A Framework for Classifying and Comparing Software Architecture Evaluation Methods', *Proceedings of the 15th Australian Software Engineering Conference (ASWEC), IEEE Computer Society*, pp. 309-319.
- [4] Basili V, Caldiera G, Rombach D. (1994), 'Goal Question Metric Approach, Encyclopedia of Software Engineering', pp. 528-532, John Wiley & Sons, Inc.

- [5] Dobrica L, Niemela E. (2002), 'A Survey on Software Architecture Analysis Methods', *IEEE Transactions on Software Engineering*, 28/7: 638-653.
- [6] Jong S, Arensbergen P, Daemen F, Meulen B and Besselaar P. (2011), 'Evaluation of research in context: an approach and two cases', *Research Evaluation*, 20/1: 61-72.
- [7] Juristo N, Morant J. (1998), 'Common framework for the evaluation process of KBS and conventional software', *Elsevier Knowledge-Based Systems*, 11/2: 145-159.
- [8] Juristo N. (1997), 'Structuring the Knowledge-Based Systems Evaluation Process', *Proceedings of Intelligent Information Systems'97*, pp. 266-270.
- [9] Sommerville I (2006), *Software Engineering*, ISBN # 9780321313799, Addison-Wesley, Wokingham, UK.
- [10] The HUMBOLDT Project, 2011, <http://www.esdi-humboldt.eu/home.html>
- [11] Alexander, I., (2005), A taxonomy of stakeholders: Human Roles in System Development, *International Journal of Technology and Human Interaction*, Vol 1, 1, 2005, p.p.23-59.
- [12] Sutcliffe, A., (2002), *User-Centred Requirements Engineering: Theory and Practice* [book], Springer.
- [13] XUnit family, toolset for Test Driven Development, URL: <http://www.martinfowler.com/bliki/Xunit.html>
- [14] Bryson J, Patton M, Bowman R, (2011), 'Working with evaluation stakeholders: A rationale, step-wise approach and toolkit', *Evaluation and Program Planning*, 34/1, pp. 1-12, ISSN 0149-7189, 10.1016/j.evalprogplan.2010.07.001.
- [15] Eden, C., & Ackermann, F. (1998). *Making strategy*. Sage Publishers.
- [16] HUMBOLDT project, 2010, Deliverable A10.6-D1 – Scenario Evaluation Report, September 2010, http://www.esdi-humboldt.eu/press/public_deliverables.html
- [17] Brooke J, SUS – A quick and dirty usability scale, In: Jordan P, Thomas B, Weerdmeester B (eds) *Usability Evaluation in Industry*, pp. 189–194, Taylor and Francis, 1996.
- [18] Finstad K, (2010), The Usability Metric for User Experience, *Interacting with Computers*, 22(2010), pp. 232-327.
- [19] Meyers G, (2002), 'the art of software testing', 2nd edition, ISBN: 978-0-471-46912-4, Wiley Publishers.
- [20] Fowler M, *Continuous Integration*, published on 01 May 2006, URL: <http://martinfowler.com/articles/continuousIntegration.html>
- [21] Code Analysis tools – Wiki - http://en.wikipedia.org/wiki/List_of_tools_for_static_code_analysis
- [22] Code Coverage – Wiki - http://en.wikipedia.org/wiki/Code_coverage

Annex-1 – Evaluation Design Template for CIM2

| Criterion ID | Main Criterion/ Concern | Sub-criterion | Evaluation Stakeholders | | | | Indicator ID | Indicator | Description/ Questionnaire | Metrics/ Answering Options | Level of Importance for Stakeholders - High (H), Medium (M), Low (L), Not Relevant (NR) | Remarks / Resources |
|----------------------|----------------------------|---------------|-------------------------|-----------|----------|--------|--------------|-----------|-------------------------------|----------------------------------|---|------------------------|
| | | | Primary | Secondary | Tertiary | Others | | | | | | |
| Perspective A | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| Perspective B | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |

Key:

Criterion ID: uniquely identifies a criterion. **Main criterion or concern:** refers to aspect to be evaluated. **Sub-criterion:** further categorisation of the main criterion. **Evaluation stakeholders:** who will participate in the evaluation of a specific aspect. **Indicator ID:** uniquely identifies associated indicator with the sub-criterion. **Indicator:** description of expected outcome. **Description or Questionnaire:** question(s) to be answered based on the assessment/evaluation. **Metrics:** indicates answering options with associated weights. **Level of importance:** indicates how important this particular aspect is for a specific stakeholder? **Remarks or resources:** indicates which resources (data, documents, services) are available to perform evaluation for this particular aspect or are there any remarks related to resource limitation and/or availability?