



CO3

Digital Disruptive Technologies to Co-create, Co-produce and Co-manage Open Public Services along with Citizens

Grant Agreement number: 822615

D4.1 Evaluation framework

Keywords

CO3project, H2020, implementation, disruptive technologies, Blockchain, Augmented Reality, Geolocation, SocialNetworking, Opinion, Formation, Gamification, Co-creation, Co-production, Co-management, Open Public Services, Social impact, Best practices

Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Co-Funded by the Horizon 2020 Framework
Programme of the European Union
CO3 Grant Agreement 822615



Authors List

Partner	Author(s)	Sections
OLA	Gregory Tsardanidis, Giannis Zgeras, Vasilis Vlahokyriakos	1-5

History

Version	Date	Reason	Revised by
0.1	11.03.2020	Realist Evaluation Description	Gregory Tsardanidis
0.2	23.03.2020	Added Pool of Evaluation Methods	Gregory Tsardanidis
0.3	20.04.2020	Added Introduction	Gregory Tsardanidis
0.4	08.05.2020	Elaborating on section 2,3	Giannis Zgeras
0.5	12.05.2020	All sections completed	Gregory Tsardanidis
1.0	19.05.2020	Version submitted to Portal	Alberto Guffanti
1.1	10.06.2020	Section 1-3 Revision	Vasilis Vlahokyriakos
1.2	20.06.2020	Added Overarching Framework	Giannis Zgeras
1.3	03.09.2020	Elaborating on section 5	Gregory Tsardanidis
1.4	05.11.2020	Ongoing work on section 5	Gregory Tsardanidis
2.0	10.05.2021	Version submitted to Portal	Alberto Guffanti

Acronyms

ACA	Augmented Commoning Area
PA	Public Administration
UCD	User-Centred Design

Authors List	1
History	1
Acronyms	1
1. Introduction	3
2. General Approach	4
2.1 Realist Evaluation	4
2.1.1 Realist Evaluation and CO3	4
2.1.2 The Realist Evaluation Stages	5
2.1.3 Formulating our theories	6
2.2 Engaging Stakeholders	9
2.2.1 Participatory Evaluation	10
2.2.2 Responsible research and ethics in service innovation	12
2.2.3 Ladder of citizen participation	13
2.2.4 Researchers as stakeholders	14
2.3 Spot the Disruptiveness	15
3. Pool of Evaluation Methods	17
3.1 Usage Data Analysis	17
3.2 Qualitative Analysis	17
3.3 Focus Groups	19
3.4 Qualitative Interviews	19
3.5 UTAUT based Questionnaires	20
4. Early Usability Evaluation	21
5. Evaluation Plan	22
5.1 Development Process	23
5.1.1 Combining Metrics	23
5.1.2 Design the evaluation with pilots	24
5.1.3 Creating the Overarching Framework	28
5.2 Overarching Conceptual Framework	29
5.3 Action Roadmap	31
Appendix	34
A. Pilot Services Reference	34
B. Usage Data	34



1. Introduction

Deliverable 4.1 aims to establish a common framework through which the results of the activities of the three pilot sites of the CO3 project will be evaluated. The framework covers the two main axes of the project: the use of disruptive technologies and the co-production of public services. These axes are combined with different evaluation aspects that are related with Working Package 5. These aspects include legal, sociocultural, economical and sustainability factors. The evaluation framework will be contextualized to fully cover the needs of the individual pilots sites. These site specific evaluation methods will be provided in D.4.2.

The evaluation process is targeted to help the pilot applications to evolve through an iterative process and to bring about useful conclusions for the project as a whole. For this purpose the general approach that we are going to follow is **Realist Evaluation**.¹ This case-based approach can be iteratively applied to provide spot-on information on the conditions that influence the implementation of the services in each pilot.

In addition, we are going to work towards a deeper engagement of stakeholders in the evaluation. For this purpose elements of the **Participatory Evaluation**² approach will be followed aiming to engage the stakeholders in the process of deciding on what it is important for them to evaluate. Moreover, we will combine elements from other methods to better understand the levels of their involvement.

Early evaluation processes will take place to ensure both the quality of the functionality provided and the service process proposed. For this purpose the Evaluation Plan will work conjointly with the Engagement Plan, as described in deliverable D.3.1. The engagement actions described there provide several opportunities to implement evaluation processes, while a number of the engagement Key Performance Indicators described could be used as basic metrics.

To transform the general evaluation framework into a detailed evaluation plan we provide a pool of evaluation methods that can be used in pilots sites. Our goal is to discuss both metrics and methods with local pilot partners and stakeholders to co-design the evaluation plan in each pilot.

Unfortunately the situation in pilot sites, due to the COVID-19 pandemic, made it extremely difficult in Spring 2020 to engage efficiently with local stakeholders. In addition, any detailed plan provided would fall short, since the condition in which the pilot services will take place had not been made clear. Thus, the first version of the deliverable focuses on our general approach.

¹ Pawson, R., & Tilley, N. (2004). *Realist Evaluation*. 2004. London: Sage.

² Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New directions for evaluation*, 1998(80), 5-23.

The second version develops our approach further and describes the specific stages the evaluation process will follow. It also presents specific action types and their data collection and analysis methods. The timetable and the ability to carry out all evaluation actions is still under review due to the ongoing COVID-19 pandemic.

2. General Approach

In this section we describe our general approach for the evaluation of the CO3 pilot services. We want to evaluate each pilot and form a coherent narrative from these evaluations, which focuses on the **use of disruptive technologies (A) in co-production of public services (B)**. These are the two primary axes of our evaluation process. All other aspects referring to WP5 activities will be incorporated accordingly in the general evaluation framework.

2.1 Realist Evaluation

The general approach we are going to follow is **Realist Evaluation**. Realist Evaluation is a case-based evaluation approach or ‘logic of enquirement’ (Pawson and Tilley, 2004)³ that originated in evaluating social programmes and can be applied to complex projects that are implemented in real social environments and involve several stakeholders.

2.1.1 Realist Evaluation and CO3

CO3 is a research project that aims to explore how disruptive technologies and the co-production concept are able to provide collaborative public services as Commons. The project aims to take under account several underlying factors to produce a broader understanding, these include legal, sociocultural and other relevant aspects. The collaborative approach of the services, as well as the variety of factors to be explored suggest that a conventional service evaluation methodology focused on effectiveness cannot be applied. Therefore we need to undertake a more explanatory approach and aim to explore what is it about this kind of intervention that works, for whom, in what circumstances, in what respects and why (Pawson et al - 2004)⁴. As such, through Realist Evaluation we perceive the CO3 project and its pilot applications as interventions, and we can evaluate which will help both to understand the social environment in which they take place and to prepare us to make recommendations for improvement that will contribute to the sustainability of the project.

In addition, the complexity of the CO3 pilot applications along with the fact that we are going to proceed to a real-life implementation in different contexts further strengthens the notion that Realist Evaluation is a suitable approach. Moreover, the CO3 pilots are going to evolve over time as stakeholders become more engaged. Realist Evaluation acknowledges and welcomes this evolution. As (Pawson et al - 2004)⁵ states it, “Interventions, however

³ Pawson, R., & Tilley, N. (2004). *Realist Evaluation*. 2004. *London: Sage*.

⁴ Pawson R, Greenhalgh T, Harvey G, Walshe K (2004) Realist synthesis: an introduction. ESRC Research Methods Programme Methods Paper 2.

⁵ *ibid.*

well planned, evolve over time as provider and recipient understanding of them increases and they begin to exert their influence over them.”

Evaluation as policy recommendation

The primary aim of Realist Evaluation is to proceed to policy recommendations. After the evaluation we will be able to provide transferable advice on the options available within a programme and the considerations (Pawson et al - 2004)⁶. Under this approach we can consider the evaluation process for the CO3 project as a formative process. The goal is to evaluate the pilot services and recommend accordingly on the future development of the project and policies affecting the underlying environment. Thus, will be able to explore the potentials of the CO3 applications and also identify environmental factors that could contribute in order for these services to have a positive impact.

As we have described in the project’s Conceptual Framework in Deliverable D1.1, we do not underestimate the potential of disruptive technologies to alter the way PA works, but we must acknowledge that the context and socio-political conditions have an impact on how technology is used and appropriated. (Papacharissi, Zizi - 2008)⁷ goes as far as to argue that “not the nature of technologies themselves, but rather, the discourse that surrounds them, that guides how these technologies are appropriated by a society.” While we do not fully follow this approach we acknowledge that it is not unlikely that new disturbing technologies could end up reproducing the same Public Administration model they aimed to disturb. In this context it is important to not only focus on the technology itself but also on the surrounding environment and the ways in which policies can affect it. Crucial for this approach is the concept of co-production since it brings an innovative approach in the relations between citizens and PA.

2.1.2 The Realist Evaluation Stages

As explained in (Hewitt et al - 2012)⁸ and initially described in (Pawson and Tilley - 1997; 2004)⁹ the Realist Evaluation has four stages:

- 1. Articulate programme theories to be tested.** In this stage we need to formulate our theories to be tested. These can be related to candidate mechanisms, contexts and outcomes

⁶ *ibid.*

⁷ Papacharissi, Z. (2008). The virtual sphere 2.0: The Internet, the public sphere, and beyond. In *Routledge handbook of Internet politics* (pp. 246-261). Routledge.

⁸ Hewitt, G., Sims, S., & Harris, R. (2012). The realist approach to evaluation research: an introduction. *International Journal of Therapy and Rehabilitation*, 19(5), 250-259.

⁹ Pawson, R., & Tilley, N. (1997). An introduction to scientific realist evaluation. *Evaluation for the 21st century: A handbook*, 405-418. , *ibid.*

2. **Collect data to test the hypotheses.** In this stage we collect data with a variety of methods: quantitative, qualitative or mixed. These methods are tailored to the theories we have formulated.
3. **Test the hypotheses.** Use the data collected to spot patterns in the outcomes of the programme evaluated.
4. **Interpretation and refinement.** Finally we try to answer our evaluation question and make policy recommendations to accordingly revise the programme.

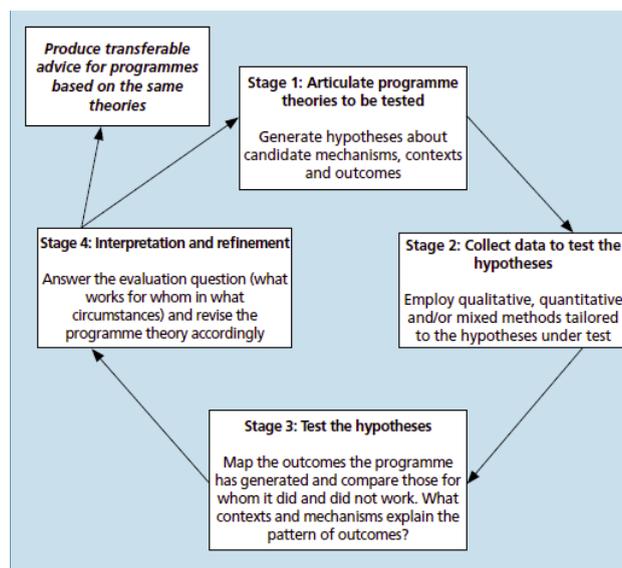


Figure from (Pawson and Tilley - 1997) as adapted by (Hewit et al - 2012)

In the case of the CO3 project we are going to work to formulate our theories (Stage 1) from the information gathered from stakeholders during the co-design phase and the project's conceptual approach as described in D.1.2. The gathering of data will take place during the pilot activities (Stage 2). The testing of our theories and the interpretation of the findings (Stage 3 and 4) will be the essential part of WP4 and will be presented in D4.3. Finally, policy recommendation will be part of WP5.

2.1.3 Formulating our theories

As we stated in the previous section the first stage of the Realist Evaluation process is to formulate a theory to be evaluated.

“The realist approach illuminates the theories that lie beneath a programme, enabling the realist researcher to examine them, identify regular patterns of CMO¹⁰ configurations and answer the realist question of ‘what is it about this programme that works for whom in what circumstances?’” (Hewit et al - 2012) ¹¹

¹⁰ This refers to Context (C), Method (M), Outcome (O)

¹¹ *ibid.*

Following this notion we need to formulate the CO3 project as a theory to be tested through Realist Evaluation. This core theory, based on the project proposal, can be summarized as following:

The capabilities afforded by disruptive technologies and the co-design process can contribute to a successful collaborative public service provision.

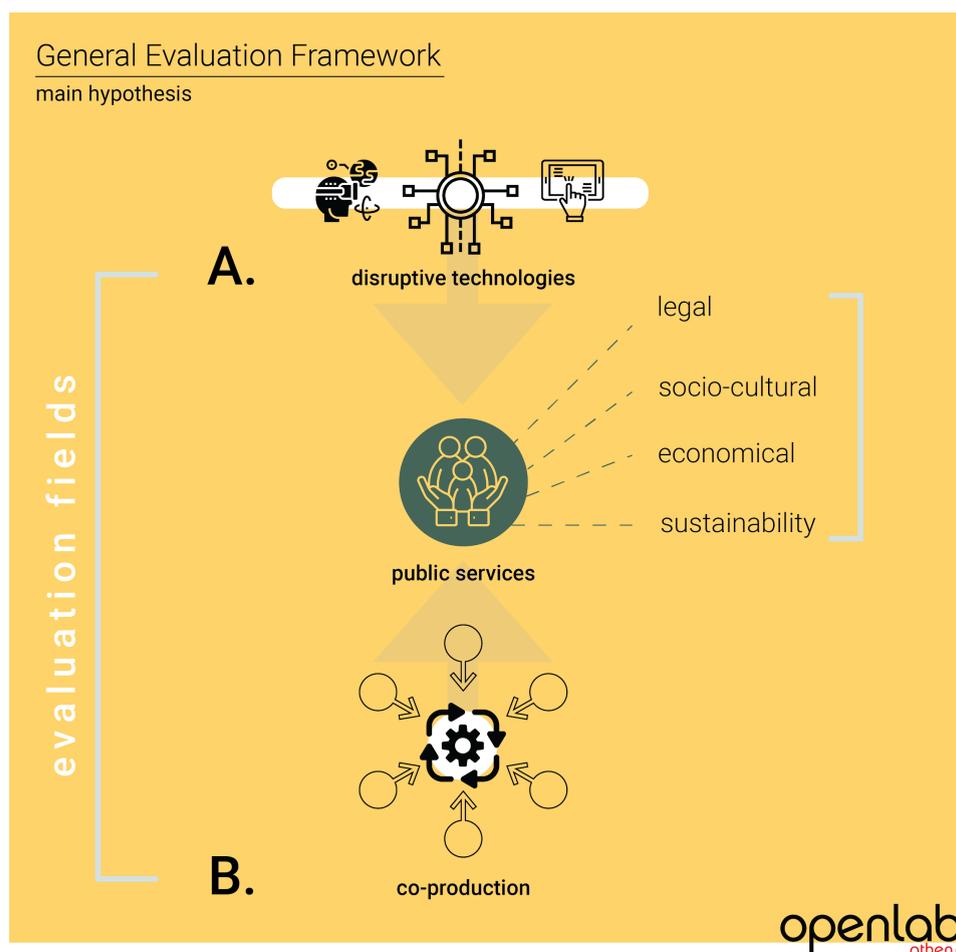
This theory summarizes the project's main goal as described in D.1.1: "The aim of this project is to explore the potential of innovating Public Services in the direction of Co-creation, Co-production and Co-management with the use of disruptive technologies". To further elaborate our theory we can state that the evaluation process wants to explore under which circumstances (Context) the use of disruptive technologies under the co-design/co-production concept (Method) ¹² can result in successful public services (Outcome).

Realist evaluations are based on an assumption that projects and programmes work under certain conditions and are influenced by the way in which different stakeholders respond to them.¹³ Restructuring the CO3 project modules to fit the above context we can define the projects/programmes to be examined as the CO3 pilots, the conditions as the factors identified for evaluation from the project proposal and WP5 (legal, economical, sociocultural, sustainability), and as the stakeholders the participants of the co-design and co-production processes including the PA, citizens, urban commoners and local businesses.

¹² Method is both co-design and the technologies designed and their affordance. In our concept disruptive technologies are also a "method" through which co-production can be achieved

¹³ Realist Evaluation, Intrac for Civil Society, 2017





CO3 main theory to be tested

This main hypothesis is the basis on which we can create hypotheses for the underlying factors we aim to explore in the following form:

The role of the [condition] environment is essential for successful collaborative public service provision by utilizing disruptive technologies and the co-design process

Where the condition we can state are:

- Legal, as of the legal environment in which the pilot operates
- Sociocultural, the cultural background of the pilot stakeholders
- Economical, the economic conditions underlying the pilot operation and the general economic environment in the pilot site

All these help us evaluate the Sustainability of the services by taking under account the resources/conditions needed on all above aspects along with the impact on social service provision.

Moreover, hypotheses can be formulated for each pilot service or group of similar pilot services. Depending on the nature of the pilot service and the perception of the

stakeholders we can make and test hypotheses that contextualizes our general hypothesis to the local conditions.

For example for pilot Paris.1 “Contributive Clinic”¹⁴ we can form a theory as follows:

The capabilities afforded by FirstLife and IRI Tools can contribute to empower and equip single mothers to tackle their children's addiction to digital displays

In addition, the concept of “successful” - as stated in our hypothesis - could also be defined at the pilot level by approaching certain aspects of social and economic impact. The variety of services provided by the CO3 project requires a thorough approach that takes under account the nature of the service, the local stakeholders and the conditions in the pilot site.

For example, for pilot Athens.1 “Groceries on hold”¹⁵ we can define impact measurements as following

- The amount of groceries raised through the service
- The sentiments of the beneficiaires
- The acceptance of the system by bench owners

All these measurements will not be combined in a strictly defined and valued cross pilot set of variables. Instead, these pilot services will work as case studies which provide insight for the circumstances under which they took place. These insights will help us discuss the ways in which our general hypothesis can become true rather than measuring if it did or not.

Further discussions took place with pilot partners and local stakeholders to fully define these pilot based theories as described in Section 5.

2.2 Engaging Stakeholders

Realist Evaluation values the role of stakeholders. As (Hewit et al - 2012)¹⁶ state: “The first aims to identify the programme mechanisms for the theoretical framework. [...] could therefore include [...] discussions with policy-makers or programme stakeholders to understand their framing of the problem and the programme’s solutions.”

We aim to deepen this process and engage proactively with the primary stakeholders and decision makers which are already on board from the co-design process. To achieve this we will combine the Realist Evaluation approach with elements from other other methods and concepts. The principles of Participatory Evaluation will guide us to better incorporate the views of the stakeholders in the evaluation process. Moreover, we will focus on how the citizens, one of our primary stakeholders, are participating in the co-production of the

¹⁴ See Appendix Table 1

¹⁵ See Appendix Table 1

¹⁶ *ibid.*

service. Finally we will aim to take a closer look at how the researchers, the partners of the CO3 project, are influencing the co-design and the co-production process.

2.2.1 Participatory Evaluation

Important part of the approach in the CO3 project is that of Participatory Action Research in order to develop community interventions that promote collaboration and commoning to wider social groups in the city. The participatory research process enables researchers to step back cognitively from familiar routines, forms of interaction, and power relationships in order to fundamentally question and rethink established interpretations of situations and strategies (Bergold & Thomas 2012)¹⁷. In this realm, our evaluation framework is important to dispute the way we used to evaluate sustainability and to acknowledge impact in social service provision, if we want to create a sense of ownership and commoning for the designed services among stakeholders and citizens.

Consequently, we will incorporate elements of the **Participatory Evaluation** approach to achieve greater involvement of the stakeholders in the evaluation process. Participatory evaluation is an evaluation strategy which implies that, when doing an evaluation, researchers, facilitators, or professional evaluators collaborate in some way with individuals, groups, or communities who have a decided stake in the program, development project, or other entity being evaluated (Cousins, Whitmore - 1998)¹⁸

With this in mind we will engage with all stakeholders of the co-design process:

- Citizens
- Public
- Commoners
- Local businesses

The intention of such an expansion on the conception of evaluators is to empower and create awareness to all stakeholders involved in designing and implementing the service and also to the people who are affected by this implementation. In particular, this increase in stakeholders participation in the evaluation process will lead: a. to **increase the utilization of evaluation results**, b. to **represent the values and concerns** of the multiple groups involved, c. to **promote the empowerment** of disenfranchised stakeholder groups previously left out of the process (Papineau & Kiely 1996)¹⁹.

What we want to achieve is understand what it is important for the stakeholder to evaluate. Our metrics will have to incorporate their interests and views. We aim also to explore how voices unheard can have a more significant role in this process. For example beneficiaries of social services that are rarely asked on how they evaluate the service provided. We are going to follow the approach of Transformative Participatory Evaluation that has as its

¹⁷ Bergold, J. & Thomas, S. (2012) Participatory Research Methods: A Methodological Approach in Motion. *Historical Social Research / Historische Sozialforschung*, Vol. 37 No. 4 (142),

¹⁸ Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New directions for evaluation*, 1998(80), 5-23.

¹⁹ Papineau, D. & Kiely, M. (1996) Participatory Evaluation in a Community Organization: Fostering Stakeholder Empowerment and Utilization. *Evaluation and Program Planning*, Vol. 19 No.1, pp 79-93

foundation principles emancipation and social justice. It seeks to empower members of community groups who are less powerful than, or are otherwise oppressed by, dominating groups. (Cousins, Whitmore - 1998)²⁰

To achieve this we are going to have discussions with local pilots and stakeholders and try to map important metrics for the underrepresented stakeholders. The co-design process has already provided us with valuable insights in this aspect and we will combine these insights with elements from the following table of participatory evaluation approaches.

APPROACHES TO PARTICIPATORY EVALUATION				
Primary Goal	Author	Stakeholder Groups	Collective Tasks	Evaluator's Tasks
Increase utilization of results	Mark and Shotland (1985) Patton (1986)	<ul style="list-style-type: none"> • decision makers • service deliverers • primary intended users • represent relevant constituencies 	<ul style="list-style-type: none"> • not specified • determine focus • methods decisions • analysis • interpretation of results 	<ul style="list-style-type: none"> • not specified • collect data • facilitate discussions
	Barrick and Cogliano (1993)	<ul style="list-style-type: none"> • nursing faculty 	<ul style="list-style-type: none"> • collect data • strategic planning 	<ul style="list-style-type: none"> • design • data analysis • training workshops • facilitate planning process
	Greene (1987, 1988a)	<ul style="list-style-type: none"> • funders • staff • board • clients • community representatives 	<ul style="list-style-type: none"> • identify purpose and audience • choose one question • select methods of inquiry • interpret findings • recommendations 	<ul style="list-style-type: none"> • data collection and analysis • structure and facilitate process • written synthesis of findings • maintain technical quality
Represent multiple interests	Mark and Shotland (1985) Shapiro (1988)	<ul style="list-style-type: none"> • decision makers • service deliverers • faculty • staff • students 	<ul style="list-style-type: none"> • not specified • answer evaluation questions 	<ul style="list-style-type: none"> • not specified • design • collect and compile data • recommend ways to improve program
Empowerment	Mark and Shotland (1985)	<ul style="list-style-type: none"> • greater power to clients and social service recipients 	<ul style="list-style-type: none"> • evaluation planning and practice • program formulation and modification 	<ul style="list-style-type: none"> • not specified
	Brunner and Guzman (1989)	<ul style="list-style-type: none"> • beneficiaries • lay facilitators • staff 	<ul style="list-style-type: none"> • design and implementation • collective analysis • report to community for action 	<ul style="list-style-type: none"> • advise and facilitate
	Whitmore (1991)	<ul style="list-style-type: none"> • community advisory committee • service users 	<ul style="list-style-type: none"> • design and implementation • act to change program • publicize results • facilitate process • report to advisory committee for advice 	

From Papineau, D. & Kiely, M. (1996) *Participatory Evaluation in a Community Organization: Fostering Stakeholder Empowerment and Utilization. Evaluation and Program Planning, Vol. 19 No.1, pp 79-93*

²⁰ Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New directions for evaluation, 1998(80)*, 5-23.

2.2.2 Responsible research and ethics in service innovation

The use of Participatory methods as a core element throughout the research program does not mean that all ethical aspects of the project are resolved because of the collaboration between researchers, stakeholders and citizens. There is a great danger, when we are discussing about “a genuinely transformative and even novel approach to governance” and service innovation, to end up to a repackaging of existing concepts to smooth the pathway for pre-defined policy goals²¹. In addition to that, it is important to be very careful with the type and the quality of the aforementioned collaboration, because of the multi-level power relations that could be produced in the pilot cases. Especially in our evaluation framework, we have the intention of collaboratively designing a process that is open to all actors and their interests, in order to abstract specific recommendations for policies related to digital transformation in public services. For example in this realm, we have to take into consideration what kind of potential disruption impacts the **concept of citizenship** by the implementation of the pilot services and which would be the **political and social consequences of transforming the nature of public services**.

Responsible Research and Innovation (RRI) is being considered as a new criterion within the EU grant submissions process in Horizon 2020, where it may become a requirement for researchers in ICT, and science and technology more broadly, to investigate the potential social impacts of proposed research²². There is a broad discussion about ethics and responsibility in research, but on all occasions these attempts of defining “what is ethical” are assuming positionality statements. In order to define what would be the “right impact” of our project, we will create a framework mainly based on three axes:

1. The potential aspect of technology in democratizing governance by promoting public participation
2. The diffusion of commons’ principles and the creation of a digital and physical space of horizontal collaboration between public authorities, formal and informal groups and institutions
3. The raising of awareness about the transformations in our digital future and the promotion of technology literacy.

These general axes about the overall process of the program could mainly function as a critical filter for our research intentions and actions. A more context specific framework is needed, if we want to avoid important misconceptions. For this reason we will take into consideration some critical issues based on literature about ethics in the field of Human-Computer-Interaction²³

²¹ Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: From science in society to science for society, with society. *Science And Public Policy*, 39(6), 751-760. doi: 10.1093/scipol/scs093

²²G. Eden, M. Jirotko and B. Stahl, "Responsible research and innovation: Critical reflection into the potential social consequences of ICT," *IEEE 7th International Conference on Research Challenges in Information Science (RCIS)*, Paris, 2013, pp. 1-12, doi: 10.1109/RCIS.2013.6577706.

²³ Brown, B., Weilenmann, A., McMillan, D., & Lampinen, A. (2020). Five Provocations for Ethical HCI Research. & Sellen, A., Rogers, Y., Harper, R., & Rodden, T. (2009). Reflecting human values in the digital age. *Communications Of The ACM*, 52(3), 58-66. doi: 10.1145/1467247.1467265

the point that so many have missed - that there are significant gradations of citizen participation²⁵. Therefore, we have to spot the level of citizens involvement in each of the main phases of the project (co-design, co-production and co-management) and highlight the different qualities and limitations emerged in every pilot. The role of citizens is a basic factor in order to assess the potential development of co-ownership of a public service in the European city and can also illuminate the power relations between the different stakeholders. The sector of public service provision is a multi-level ecology of interconnected economic and political interests, which will be transformed by citizens participation and control.

Consequently, the ladder of participation has a double function in our evaluation framework:

- Demonstrate how the level of citizens adaptation in participatory processes reflect on project's key ideas
- Inform us about the willingness of municipalities and main stakeholders to overcome legal barriers and share parts of their authority in a more decentralized logic

Finally, another crucial point of the “ladder” is also the evaluation of the participatory aspect of the applied methodology. By qualitatively acknowledging the level of citizens participation in every phase of the project, we could draw meaningful conclusions and reflections about the openness of the process itself which will lead to more comprehensive policy recommendations.

2.2.4 Researchers as stakeholders

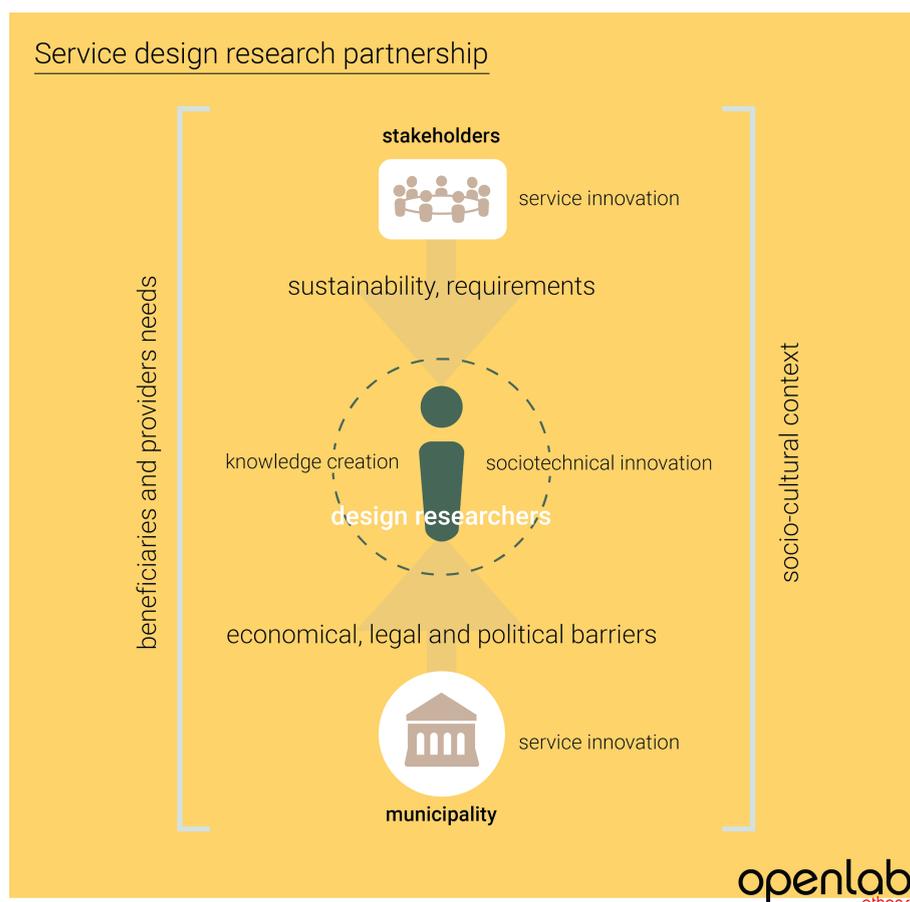
The whole project of CO3 is based on multi-level research assumptions made beforehand and influenced by personal values of researchers. The idea of a more socially sensitive framework for technology by involving citizens, commoners and public authorities had been affected by the different value conceptions of stakeholders and the chaotic dynamics among them. Thus, the role of the design researchers becomes central not only through the co-design process, but more importantly in the co-production and co-management of the designed service. The relationship between stakeholders, municipalities and citizens generates many tensions and points of conflicting interests throughout the project that should be resolved by researchers' interventions. Therefore, *that tension lies at the core of the interconnection. It is the complexity of this interconnection, and in the tension between its different qualities, that imaginative and innovative design work takes place* ²⁶. But also, this situation changes the focus of some aspects of the project and it is crucial to map those displacements in personal values and positions of researchers during the process. To sum up, the researchers by facilitating and guiding the project have to cope on the one hand with the challenging condition of creating “common ground” among partners and groups with distinct value conceptions, and on the other hand with their personal pre-conceptions of the project plan.

²⁵ Arnstein, Sherry R.(1969) 'A Ladder Of Citizen Participation', Journal of the American Planning Association, 35: 4, 216 – 224

²⁶ Harold G. Nelson and Erik Stolterman (2012) *The Design Way: Intentional Change in an Unpredictable World*. The MIT Press.

Consequently, in terms of the evaluation framework it is important to observe and report:

- Practices and methods that helped to resolve tensions and conflicts among the stakeholders, and if these practices changed (and to what extent) changed the orientation of the service
- Differentiations and displacements in researchers' views throughout the process.
- Relevance between the preconceptions of value and the final outcomes of the project ecology



CO3 research ecology of interconnections. Based on the model of service design research partnership from Nelson and Stolterman ²⁷

2.3 Spot the Disruptiveness

One of the key aspects of this evaluation process is the disruptiveness. We want to see if both the technology (A) and the co-production (B) are able to disrupt the more traditional / transactional public service delivery process. This notion also broadens the scope of the research on citizen participation described in 2.2.3. since the quality of participation achieved is part of the disruption that the technologies and the methodology has caused.

²⁷ *ibid.*

For this purpose we are going to work with local partners to see how the pilot services are linked with existing or similar services:

- In the case of Turin, for example, we can see changes in measurements regarding the CdQ. The CdQ social centers have processes to evaluate their activities and we can use the data they collect to make comparisons with the new co-produced pilot service ²⁸
- In the case of Athens.2 “Urban Modelling” we could make comparisons with standard procedures of formal discussion between city stakeholders. ²⁹

Moreover, we can examine the measurement methods already applied in existing similar services in the pilot sites and decide whether to include them in our evaluation plan in order to have comparable data where possible.

The system and the way it is designed disrupts the legal, socio-cultural and economical aspects. We aim to measure whether these technologies are disrupting the more transactional (or non-collaborative) way these municipalities are used to deliver these services. If these technologies and pilot services do disrupt the way municipalities are used to delivering these services, we then want to find out why, how, for whom, in which circumstances (legal, sociocultural etc).

Axes for evaluating the disruptiveness:

- a. Potential **displacement level of incumbent processes** in the delivery of social services through the pilot case
- b. **Unexpected conditions and capabilities** created. Potential emergence of new social interactions through the designed services.
- c. The relation between **performance, cost, reliability and adoption rate**, as indicators in comparison with past social services.
- d. **Impact on existing structures** of municipalities and stakeholders
- e. **Convergence of multiple** stakeholders and citizens groups, who did not take part in the first phases of the project

Based on the aforementioned attributes we could spot the disruptive influence of proposed technologies by asking in each case scenario whether there were any of the following³⁰:

- Delivers a capability at a previously unavailable level, which may create disruptive force
- Combines with other technologies to create synergies, which may also be disruptive
- Generates services with new performance attributes that may not previously have been valued by existing end users
- Requires users to significantly change their behavior to take advantage of it
- Changes the usual service and technology paradigms to offer a competitive edge

²⁸ See Appendix

²⁹ See Appendix

³⁰ National Research Council. 2010. *Persistent Forecasting of Disruptive Technologies*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/12557> p.33-47)

- Exponentially improves the value received by the users

3. Pool of Evaluation Methods

In the following section we describe a number of Evaluation Methods that we can use in the pilot specific evaluation activities.

3.1 Usage Data Analysis

Usage Data provided by the CO3 application can be analysed in a quantitative way. The usage data can include different types of data depending on the case and the CO3 functionalities used. These can be:

- General Data: User Registered, Visits in the App, Time spent.
- AR Functionality: Items Placed, Items Views, ACA visited, Items Interactions, Type of Items placed.
- Wallet Functionality: Tokens created, Tokens received, Tokens consumed, Transaction Volume, Token withhold time
- FirstLife Functionality: Places mapped, Places clicked, Comments Made, Type of Places
- LiquidFeedback Functionality: Votings held, Votes casted, Comments Made, Proposals Made
- Gamification Layer: Points Earned, Levels passed, Time to pass level

These data could also be examined along different dimensions such as:

- User Role
- User Physical Area
- User ACA
- Sex
- Age Group

These Usage Data, depending on their form could be analyzed with different methods. For example Discourse Analysis could be used to analyze the discussions made in First Life and Liquidfeedback. In other cases such as the Wallet Tokens we can use the data collected as a method for interviewing users about their involvement in this scenario and thus understand their perspective. For example, we can ask users with high Token consumption which features (e.g. Gamification) drive this behavior.

3.2 Qualitative Analysis

In order to create an overall image about the impact of pilot services in the field, we will use a mixed method approach on analysing data collected from interviews, focus groups and more importantly on comments made in CO3 platform. We will use thematic analysis on transcriptions from interviews and focus groups with a view to develop a broader framework of concepts that will emerge from participants' conceptions and experience of pilot

services, while on comments made in Liquid Feedback and First Life, we will exploit discourse analysis for the purpose of investigating predefined variables.

Thematic analysis

Thematic analysis is a method for identifying, analysing and reporting patterns(themes) within data³¹. Our intention is to gain a deep understanding on all possible aspects of the designed services without being theoretically bounded. In this respect, participants and users during the implementation of the services, will have the opportunity to evaluate the experience of the service and express their views through semi-structured interviews or focus groups and workshops. We will analyze these data based on the central axes of our evaluation framework (legal, economic, sociocultural, sustainability), while in parallel we will spot new themes emerging that have a significant concentration in our data. In particular, we will follow the 6 phases proposed by Virginia Clarke and Victoria Clarke:

1. **Familiarizing yourself with your data:** Transcribing data (if necessary), reading and re-reading the data, noting down initial ideas.
2. **Generating initial codes:** Coding interesting features of the data in a systematic fashion across the entire data set, collating data relevant to each code.
3. **Searching for themes:** Collating codes into potential themes, gathering all data relevant to each potential theme.
4. **Reviewing themes:** Checking if the themes work in relation to the coded extracts (Level 1) and the entire data set (Level 2), generating a thematic 'map' of the analysis.
5. **Defining and naming themes:** Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells, generating clear definitions and names for each theme.
6. **Producing the report:** The final opportunity for analysis. Selection of vivid, compelling extract examples, final analysis of selected extracts, relating back of the analysis to their search question and literature, producing a scholarly report of the analysis.³²

Discourse analysis

Discourse analysis is a research method for studying and analyzing in depth written or spoken language in relation to its social context. Social media now provide a site of fundamental shifts in communicative practices and help us to understand the ways in which different cultural communities interact in terms of concepts, values, rules, strategies, means, channels, purposes and consequences (Gwen Bouvier - 2015).³³

In the context of CO3 pilots, comments made in FistLife and LiquidFeedback can be used to reflect the way users conceived the public services and municipalism in each case

³¹ Virginia Braun & Victoria Clarke (2006) Using thematic analysis in psychology, *Qualitative Research in Psychology*, 3:2, 77-101

³² Virginia Braun & Victoria Clarke (2006) Using thematic analysis in psychology, *Qualitative Research in Psychology*, 3:2, 77-101

³³ Gwen Bouvier (2015) What is a discourse approach to Twitter, Facebook, YouTube and other social media: connecting with other academic fields?, *Journal of Multicultural Discourses*, 10:2, 149-162

scenario. This material is a way to project and analyze how disruptive technologies impact on the citizens' involvement in governance and decision making. Thereby, the process of profiling users statements, categorizing and connecting them in time, could answer assumptions about the disruptiveness of these technologies on the idea of citizenship and participation.

Finally, the aspect of co-production through this data analysis, especially in FirstLife and LiquidFeedback, could point out useful outcomes about how technology could function (positively or negatively) in terms of direct democracy and self-governance. These different research inquiries will be discussed with stakeholders in every case scenario in order to be adapted to special characteristics and requirements of every pilot.

3.3 Focus Groups

Focus groups are a form of group interview, which work as a research or evaluation tool to gather information. Every pilot should organize focus groups in which they will include stakeholders who took part in the co-design workshops, researchers and end users of the services. Listening and observing interactions can help us to appreciate how people think and feel about each service and the interconnections produced between the different stakeholders in co-design, co-production and co-management.

Focus groups specifications:

- composed of 5 to 8 people
- 1 or 2 facilitators
- small presentation of service delivering and progress
- clearly defined topics (such as users experience, feelings, functionality of interface, ethics etc.) and questions
- duration: 60-90min
- recordings of the meetings

An important step in preparing for the focus group interview is the development of topics and of a set of questions that encourage participants to respond, and solicit the information needed from the group. This preparation would be connected with the implications from the aforementioned Usage Data Analysis and Discourse Analysis

3.4 Qualitative Interviews

Qualitative interviews is a way of evaluating services that are aimed at individualized outcomes, and in parallel define the differences between participants' experiences. We propose semi-structured interviews with stakeholders, producers and beneficiaries of the service. In order to facilitate the meetings, it is important for every pilot to create an "interview guide" adapted:

- to the qualities of each case scenario,
- to the differences of every target group,

- to compare current conditions created by the project with past services of social provision.

This process will give the participants the opportunity to describe what was important to them using their own words rather than being restricted to predetermined categories of a questionnaire. For example, for pilot Athens.1 “Groceries on hold”, through interviews with beneficiaries we can discuss themes related to “social stigma” for those who are in need of food and if the use of social wallet had a positive impact on this aspect. On the other hand, interviewing bench owners, could inform us about the level of their engagement with the technicalities of the project and illuminate us about possible problems in usability of the service. It is crucial for interviewers to give space and time to participants to express and develop their thoughts, in order to explore unexpected themes or ideas raised by them.

3.5 UTAUT based Questionnaires

Based upon conceptual and empirical similarities across prior technology adoption models, (Venkatesh et al. 2003)³⁴ developed the Unified Theory of Acceptance and Use of Technology (UTAUT) that not only underscores the main individual-level factors that affect technology acceptance, but also identifies the contingencies that would amplify or constrain the effects of these factors.

Performance expectancy is defined as the degree to which an individual believes that using the system will help him or her better attain significant rewards. Performance expectancy was found to be a significant determinant of behavioral intention, with its effect varying across gender and age such that the effect is strongest for younger men.

Effort expectancy is defined as the degree of ease associated with the use of the system. It was found that the effect of effort expectancy on behavioral intention varies across gender and age such that the effect is strongest for older women in early stages of experience.

Social influence is defined as the degree to which an individual perceives as important that others believe he or she should use the new system. The effect of social influence on behavioral intention was found to be contingent on gender, age, experience, and voluntariness, such that it is the strongest for older women in early stages of experience in mandatory contexts.

Facilitating conditions is defined as the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system. It was found that the effect of facilitating conditions on technology use was moderated by age and experience such that the effect was strongest for older workers in later stages of experience.

³⁴ Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.

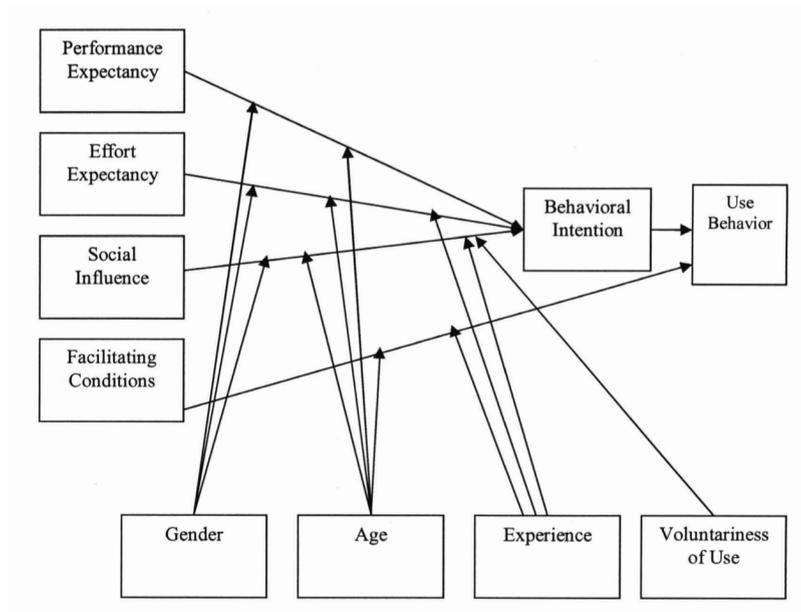


Figure from Venkatesh et al. (2003)

4. Early Usability Evaluation

We aim to undergo early usability evaluation in conjunction with the Quality Assurance processes taking place in WP2. Early usability testing can help eliminate factors that could affect the services and are not linked with the axes we aim to evaluate but merely with User Experience issues.

For this purpose we will undergo tests with potential users and stakeholders to evaluate the usability of the system as well as with expert evaluators. We have been testing every release of the application to ensure usability and we are in the process of organizing a focused usability workshop with the help of professors from the Informatics Department of the University of Athens.

Our general guidelines will be the following the Heuristic Evaluation approach (Nielsen, Jakob 1994)³⁵:

1. **Visibility of system status:**

The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

2. **Match between system and the real world:**

The system should speak the user's language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.

3. **User control and freedom:**

Users often choose system functions by mistake and will need a clearly marked

³⁵ Nielsen, Jakob (1994). Usability Engineering. San Diego: Academic Press. pp. 115–148. ISBN 0-12-518406-9.

"emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.

4. **Consistency and standards:**

Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.

5. **Error prevention:**

Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.

6. **Recognition rather than recall:**

Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

7. **Flexibility and efficiency of use:**

Accelerators—unseen by the novice user—may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

8. **Aesthetic and minimalist design:**

Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

9. **Help users recognize, diagnose, and recover from errors:**

Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

10. **Help and documentation:**

Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

5. Evaluation Plan

In the following section we will describe the general Evaluation Plan we developed. The plan follows the conceptual approach described and was developed after extensive discussions with local pilot partners. First we analyze the process through which the evaluation plan was developed and then we go on to describe the phases of the evaluation plan and the specific action types. The contextualized evaluation actions will be developed as part of Task 4.2 and will be part of D.4.2.

In general, we started from the Conceptual Approach of Realist Evaluation and worked with the pilots on the potential site-specific actions. Then by having the input from the pilots we went back to form the Overarching Evaluation Framework. With that in mind we re-organized the evaluation actions in separate phases and developed a coherent process to be carried out in practice.



5.1 Development Process

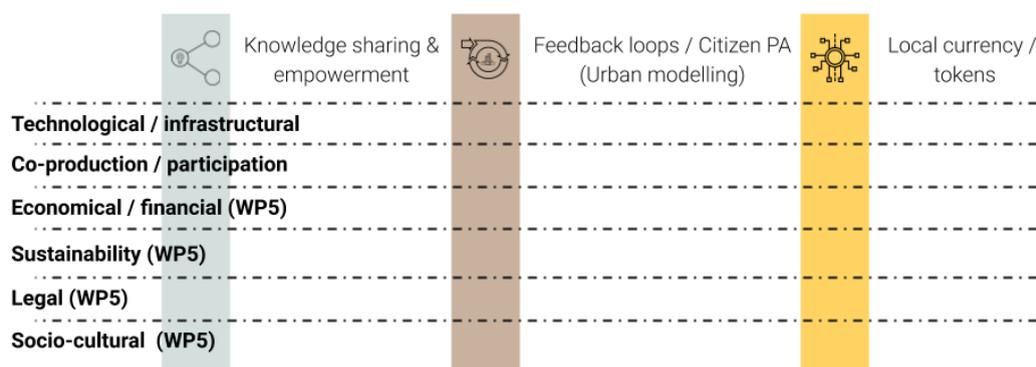
5.1.1 Combining Metrics

Our aim was to combine metrics from different perspectives fit to the type of the co-produced services. This Matrix of evaluation methods facilitated the overall analysis of the project. Mapping the methods produced by the Realist Evaluation process helped us understand the scope of the services and the metrics collected.

To be better unify our approach we have categorized the pilot services as following:

- **Knowledge sharing & empowerment.** Here we group services that focus on providing knowledge to communities. Here we include services Paris.1 Contributive Clinic ³⁶ and Paris.3 Blockchain Knowledge Registry ³⁷
- **Feedback loops / Citizen PA (Urban modelling)** Here we group services that aim to allow citizens to map their views in the urban space. Here we include services Paris.2 Digital urbanism in Augmented Schools ³⁸ and Athens.2 Mapping community needs on empty buildings. ³⁹
- **Local currency / tokens** Here we group services that use tokens to distribute value in local communities. Here we include the elements of casa del Quartiere Tokens of the Augmented Commoning pilot in Turin ⁴⁰ and Athens.1 Turin Groceries on hold ⁴¹.

Of course all these services combine different elements and cannot be completely described in this context. Nonetheless this level of abstraction can prove useful when aiming at making conclusive observations for the project as a whole.



Matrix Combining Metrics from Services and Aspects

³⁶ See Appendix A. Pilot Services Reference

³⁷ a.p.

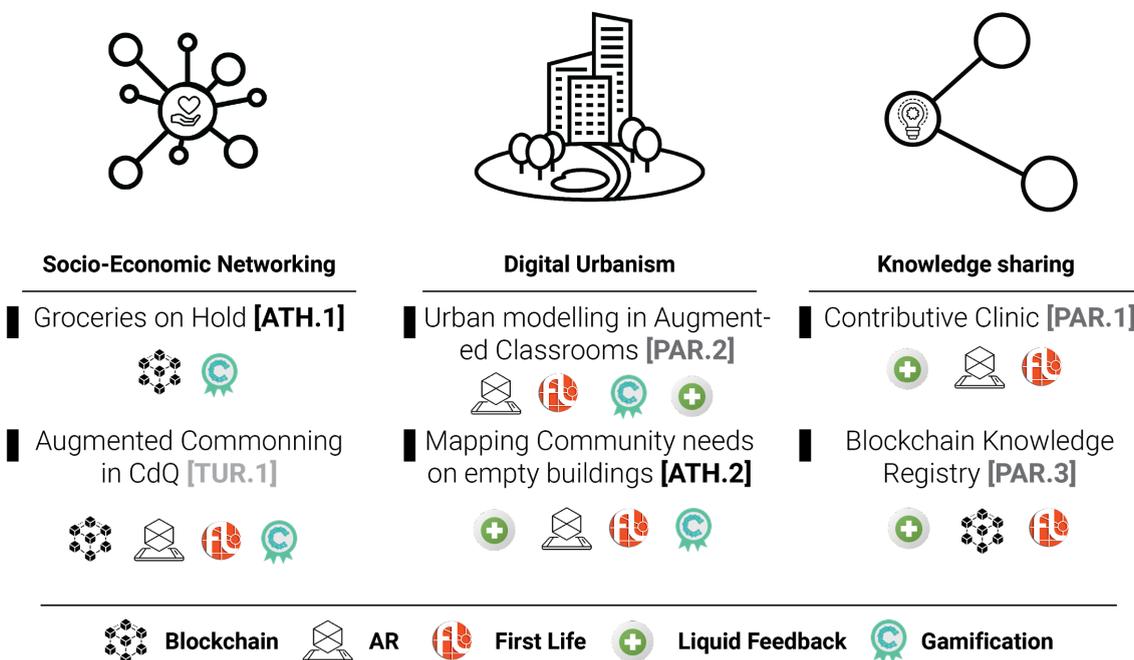
³⁸ a.p.

³⁹ a.p.

⁴⁰ a.p.

⁴¹ a.p.

The process of CO3 during the first year, generated broad variations relating to: the “nature” of the designed services, the use of technology made and the sectors of public administration in which these services will be applied. In this case, the collaborative approach of the services, as well as the variety of factors to be explored suggest a primary categorization of the pilot cases, in order to form an overarching framework that would be capable of creating useful conclusions for future policies relating to digital transformation. In this realm we end up in three different categories (socio-economic networking, digital urbanism and knowledge sharing) which cover in detail the aspects of 6 services and enable a cross-pilot evaluation of disruptive technologies in a variety of conditions and uses.



*Presentation of Pilot Cases.
Their categorization and the Technologies they use*

5.1.2 Design the evaluation with pilots

To develop the evaluation plan we engaged extensively with local pilots. We organized online meetings in May and June 2020 in which we discussed with the pilots the hypothesis they want to test and the specific actions they will carry out.

This process, of developing the evaluation process along with the pilots, was crucial because it allowed us to have a contextualized perspective on the services and their

impact. This helped us form, in a specific way, the “theory to be tested” as required in Realist Evaluation ⁴²

In addition, the discussion with the pilots incorporated the feedback from local stakeholders with which the pilot partners are deeply engaged. This helped us form an evaluation process in which we take under account the aspects stakeholders deem important to be evaluated. This way we incorporated elements of Participatory Evaluation ⁴³. Our aim for more direct engagement could not be fulfilled due to the circumstances created by the Covid19 pandemic. Our evaluation process is open to alteration during the pilot implementation when we will hopefully have opportunities to engage fruitfully with local stakeholders.

Pilots have also an extensive knowledge on the public service provision in their area and the underlying circumstances. Keeping them on board allowed us to be able to spot similar services that can be “disrupted” ⁴⁴ and design realistic evaluation actions that can be carried out meaningfully during the implementation.

Finally during the meetings with the pilots we took under account the engagement plan ⁴⁵. This was very helpful for designing the evaluation actions. We were able to use the user engagement points defined in the engagement plan to place evaluation actions. In some cases such as Athens.2 ⁴⁶ engagement actions were also the primary evaluation actions.

Forming Hypothesis & Sub-Hypothesis

At first we worked with the pilot partners to create a main hypothesis for each pilot case. This hypothesis represents the pilot case as a whole. For example the main hypothesis for Paris.2 ⁴⁷ is

The enhancing of the capabilities of the students, professors and professionals via CO3 AR App and FirstLife along with the use of Minetest can contribute to the understanding of the new ways in which inhabitants can be put inside the loop of deciding for urban planning construction.

We follow the structure of the main project hypothesis described in 2.1.3. We specify further the actors, the technology to be used and the goal of the service. Actors here are students, professors and professionals. Technologies are AR App, FirstLife and Minetest and the goal is to engage inhabitants of the area in the process of urban planning. Technologies that are not part of the CO3 project but are used in the service, in this case Minetest, are also included in the hypothesis since they also contribute to the disruptiveness of the service.

⁴² See Sections 2.1.2 and 2.1.3

⁴³ See Section 2.2.1

⁴⁴ See section 2.3

⁴⁵ See D.3.1

⁴⁶ See Appendix A. Pilot Services Reference

⁴⁷ a.p.

To better enrich our approach we went on to define specific sub-hypothesis for each service's main hypothesis. These sub-hypothesis aimed to enlight specific aspects of the pilot service, important to the pilot organization or local stakeholders. For example a sub-hypothesis for Paris.2 pilot is:

Building information management (BIM) technologies articulated with Minetest and CO3 platform will help students to understand the problems and the potentials opened up by digital platforms contributing in the urban environment

In this case we have a particular focus on students and how their understanding changes when they use the digital tools provided. This particular sub-hypothesis can lead us to create an evaluation action, such as a questionnaire to be given before and after the pilot implementation, to explore how the views of the students on the impact of digital platforms on the urban environment have swifited.

In another example from Paris.2 we have the following sub-hypothesis:

The use of the system will lead to the development of links between the educational institutions and professional paths in the territory

This sub-hypothesis reflects the needs of local stakeholders who are greatly concerned with employment opportunities of students. Nonetheless, it is not always possible to examine all hypotheses provided. During the design of the contextualized evaluation actions these sub-hypothesis will be re-examined and prioritized according to the situation on the field and the proximity to the project goals.

Disruptive Hypothesis

Even though most hypotheses have an element of disruptiveness we aimed to develop specific *disruptive hypotheses* that aim to map existing services, similar to pilot ones, and spot the disruptiveness. This way we are able to better understand the context in which the pilot service will operate and what existing provisions will disturb.

For example in Athens.1 we have the following disruptive hypotheses:

Hypothesis	Current State
Citizens/donors will be eager to contribute economically with cash	Citizens/donors contribute with dry food to social initiatives (e.g. Allos Anthropos, Mirmigi, Steps)
Citizens in need prefer to be able to shop with tokens in the flea market than get dry food for free	They get groceries from KYADA and Social Initiatives (e.g. theCupboard, Mimigi)
Producers are more eager to conceptate the digital tokens from the wallet then the paper based food stamps	The Prefecture of Attica and the Union of Sellers provide paper stamps to buy groceries in flea market to citizens in need

The city prefers to support the co-produced food provision by citizens than run services supported by private company donors	KYADA food provision is mostly funded by private company donations
--	--

In the first column of the table we see the hypothesis we make for the implementation of the pilot service and in the second column we see the current state of service provision. We aim to draw direct links between the pilot service and the operating service in order to understand the effect the pilot service can have in the surrounding social and economic environment. This is also of great importance for the pilot partners since they aim to have a tangible way to understand change.

Mapping Evaluation Actions

On the final stage of the discussion with the pilots we went on to map the evaluation actions that will be carried out in every pilot. These actions are not final, they will be finalized during Task 4.2.

The goal here was to map these evaluation actions in order to:

- Design actions able to collect data in order to test the hypotheses developed
- Attempt to link evaluation actions with engagement actions to achieve a more coherent pilot development course of action
- Use them as feedback to create the overarching framework and decide on the general action types and phases

In the case of Athens.1 for example we have action *A.1.2 “Ask citizens why they donated (or why not)”* that will be carried with *Short Interviews & Questionnaire* and is aimed to collect data to test sub-hypothesis *A1.SH1 “Citizens will be eager to contribute economically to the on-hold food provision service.”* In addition this evaluation action was linked with engagement action *B.4 “Presence in market to get supporters“* which will be carried out in the market during pilot implementation. We see here that an evaluation action is designed, perfectly fitting the engagement plan, to help us understand how one of our major users, the citizen/donor, perceived the system.

In another example from Turin the action *T.2.4. “Measure impact of the service on the work of CdQ managers and organisers”* was proposed to test sub-hypothesis *T2.SH3 “The system improves daily work and does not create a bigger administrative workload”*. This is an important measurement for local stakeholders and for the sustainability of the service. It was concluded that a separate *Focus Group* workshop should be organized and carried out to understand how the CdQ managers and organiser perceive the pilot service in terms of workload.

With this approach we were able to go into detail and determine with the pilot how we can proceed to cover each hypothesis. Valuable insight was provided which helped us understand how to form the more general action types, phases, variables and analysis methods that are described in the next sections. We started from a general conceptual approach, then focused on each pilot and with all these feedback we went on to build a coherent overarching framework that covers pilot needs.

Finally, designing the actual evaluation actions with the pilots was an important reality check. We were able to understand the environment in which pilot partners operate, their abilities and limitations and design actions that can be realistically carried out within the scope of the project.

5.1.3 Creating the Overarching Framework

The process of designing our evaluation strategy along with pilots was based on a mixed method approach of Realist and Participatory evaluation. In this respect, the pre-defined concepts of the program, such as disruptiveness or digital transformation of social services, formed different sets of hypotheses and sub hypotheses for each case scenario. The meetings with pilots generated a broad variety of themes which can be organized in the aforementioned 3 basic categories of Socio-Economic Networking, Digital Urbanism and Knowledge Sharing. These three axes are our main orientation for the Conceptual Framework of the Evaluation Plan, which also functions on a cross-pilot level in order to acknowledge the way pre-defined concepts appear during the implementation of the services. In addition to these, themes which may arise in pilots throughout the data collection phase will be combined to the overarching framework, where parameters will be evaluated and interconnected with purpose to conclude to several generalizations about technology and service design.

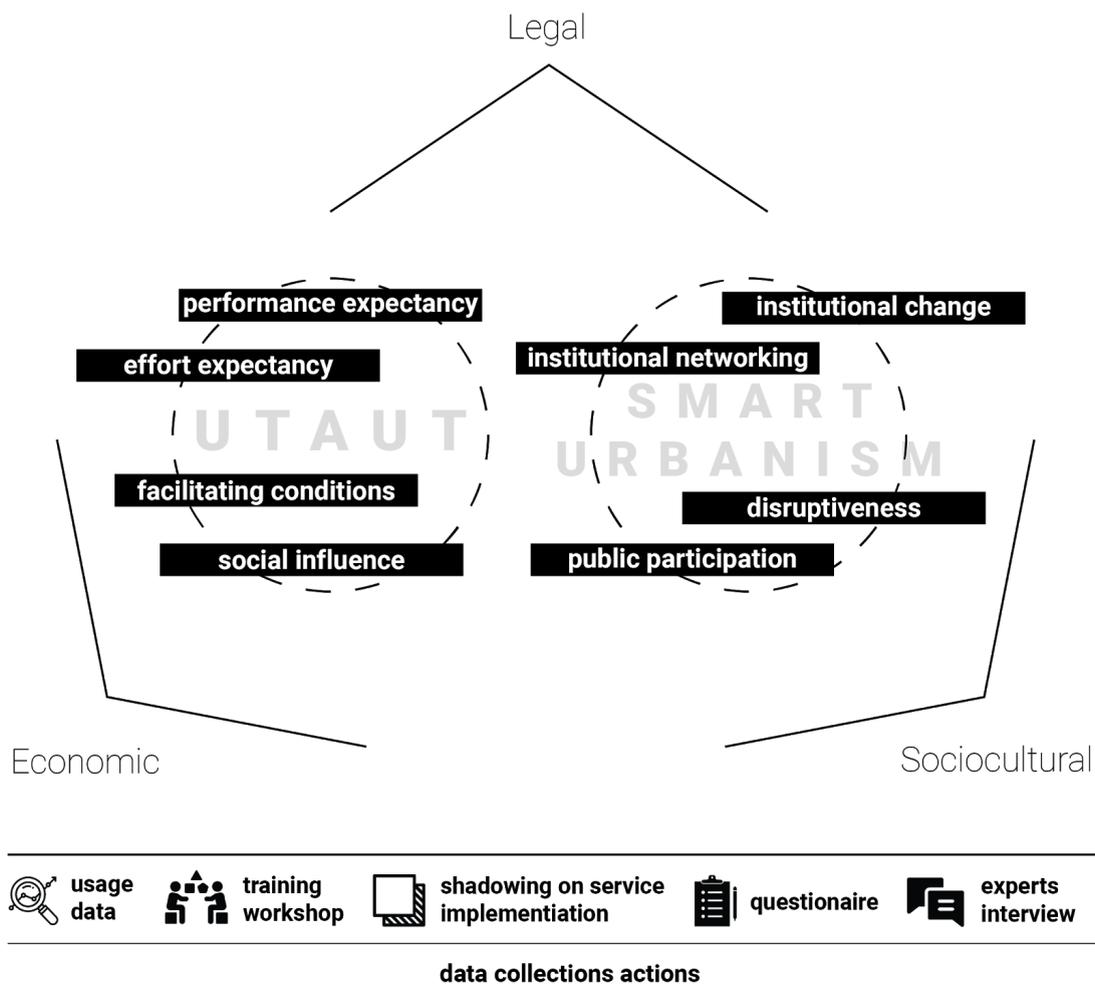
The Conceptual Framework was aligned with quantitative and qualitative research methods of data collection and analysis. We specified fields of interest which are significant for each pilot and adapted our methodology. This process remains open in order to map differentiations from past conditions in the field or transformations during the service implementation.

Defining Variables/Concept

This general framework and actions were developed in order to help us identify measurable points in each service. For this purpose, we used different variables or concepts for each aspect of the project in order to assess the quantitative data collected from users behaviour in the CO3 platform but also to integrate in the evaluation framework qualitative implications relating to socio-cultural and legal factors. In particular these variables will let us quantify qualitative aspects of the services like **disruption** or **institutional change**. Based upon the meetings with local partners who will implement the pilot services, we abstracted parameters or fields to be evaluated that are in common between different case scenarios. In parallel, these primary variables will be enriched with other themes that will emerge from field interviews and focus groups with participants and users of the pilot services, which will create the lenses of thematic analysis on data collected.



5.2 Overarching Conceptual Framework



Indicative diagram of how primary variables reflect on project's aspects and how could be measured and evaluated through data collection actions

As we described above we developed this Overarching Evaluation Framework through the meetings we had with local pilot partners. By analyzing the Hypothesis formed in that process we are able to create a more abstract and generalized conceptual framework through which we will observe the data collected and analyzed in order to draw conclusions. Even though we have formed hypotheses, following the Realist Evaluation method, we must take under account that our goal is not to test these hypotheses as true

or false but rather to understand “what works for whom and under what circumstances”.⁴⁸ In this perspective it is very important to report failures and shortcoming as well as to analyze the environment in which these pilot interventions will take place.

In the same spirit the concepts described in this framework should be more considered as conceptual lenses rather than specific variables. The complexity of the evaluation process, 3 pilot cities with different services and technologies used, as well as the variety of aspects we need to explore (economic, legal, sociocultural, sustainability) leads us to take a step back and aim to understand the big picture through the observations made. As such we will undertake a general *inductive approach*. We do not expect to draw conclusions from a specific metric or evaluation action. We aspire that all these different observation and evaluation actions when combined will allow us to better understand the impact of our pilot services and technologies. Nonetheless, an underlying theoretical framework is needed to be used as a starting point for our further analysis.

This framework has two parts. The first part focuses on the technologies and is inspired by the Unified Theory of Acceptance and Use of Technology (UTAUT)⁴⁹ The four concept inspired by the UTAUT variable are⁵⁰:

- Performance Expectancy, defined as the degree to which an individual believes that using the system will help him or her better attain significant rewards.
- Effort Expectancy, is defined as the degree of ease associated with the use of the system
- Social Influence, is defined as the degree to which an individual perceives as important that others believe he or she should use the new system.
- Facilitating Conditions, is defined as the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system.

Even though we are not oriented to using the UTAUT methodology as the structured quantitative method it is, we understand that it's basic variables summarize successfully several of the concepts emerged through the discussion with pilots. It also aligns with the goals set by the proposal. Apart from the clear need to evaluate the use of disruptive technologies as such we also observe a link with other dimensions set by the proposal. An example of this is the *Facilitating Conditions* of the UTAUT methodology with the Legal aspect set by the proposal. Another example is the link of the *Social Influence* with the sociocultural aspect, while *Effort Expectancy* can be partly aligned with the requirement for usability evaluation. We aspire that a relatively broad interpretation of these UTAUT variables will prove useful in both designing and analyzing the evaluation actions.

In the second part we focus more on the Service evaluation and try to collect different concepts with the scope of *Smart Urbanism*, in order to investigate and acknowledge possible new sociotechnical relations emerged throughout the implementation of the designed services. Smart Urbanism is argued that “provides flexible and responsive means of addressing the challenges of urban growth and renewal, responding to climate change,

⁴⁸ Pawson and Tilley - 1997

⁴⁹ See Section 3.5

⁵⁰ Venkatesh et al. 2003

increasing resilience, promoting sustainable economic growth and building a more socially inclusive society⁵¹. According to the CO3 project, our intention is to examine how our pilot services will impact urban spheres related to:

- a. Digital urbanism
- b. Socio-economic networking
- c. Knowledge sharing

These three broad categories of our proposed services have to be analyzed deeper in order to provide some indicators that could spot possible changes, impacts or failures in the ecosystems of the three cities. Based on the the meetings we had with local pilot partners and the hypotheses formed we ended up with 4 variables inspired by the concept of Smart Urbanism⁵²:

- **Institutional change** refers to transformations of processes, protocols or management models of public sector or other institutions involved in the delivery of the service
- **Institutional networking** is defined as relations emerged between formal or informal organizations during the implementation of the service in order to deliver the service or to transform conditions of existing ecosystem
- **Disruptiveness** refers to practices, workflows, behaviors or views which transformed because of the experience gained during service implementation
- **Public participation** is defined as the degree in which the proposed services promoted civic engagement and created opportunities of interaction between citizens and public administration

These abstract theoretical concepts will help us organize and analyze the collected data on a more qualitative level in order to identify the organizational and social implications of disruptive technologies in the field. Our intention is to have a more holistic approach on how lessons-learned throughout the service implementation could provide useful insights and conclusions in the aspect of policy recommendation. Moreover, it is also important to map the way the three different cities complicate, translate, enable or resist the CO3 principles and the digital transformation of public services through disruptive technologies.

5.3 Action Roadmap

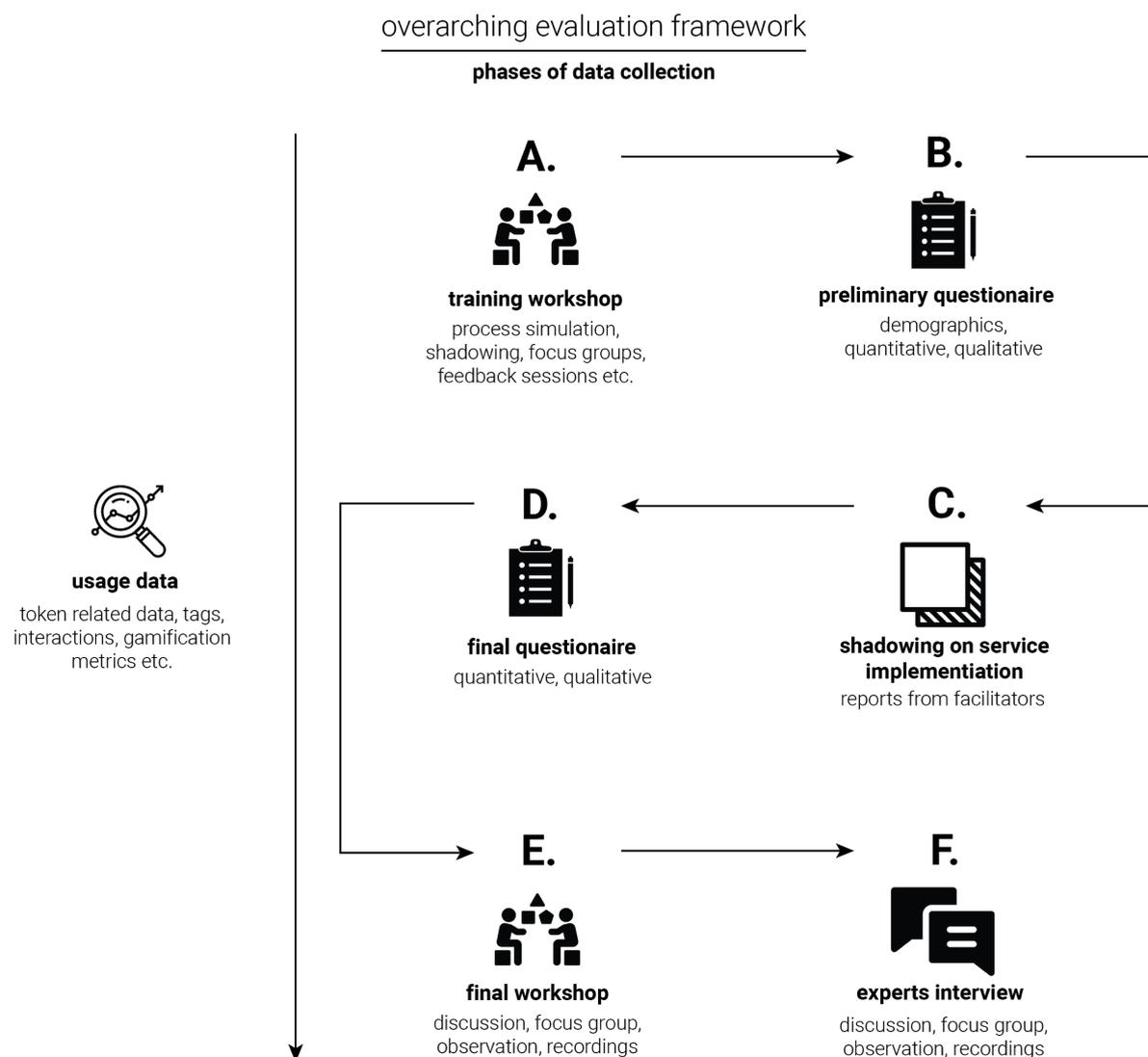
Based on the meeting with the pilots and the pool of evaluation methods we went on to combine the different needs and produce a consistent roadmap for the development of the evaluation plan in each pilot. This roadmap aims to provide local pilots and partners in Task 4.2 a clear way to organize the contextualized evaluation actions. This roadmap is aligned as much as possible with the engagement plan in order to make it easier to implement. Not all phases must be implemented in each pilot in the same way and with the same order. It

⁵¹ European Commission (2013a) Energy technologies and innovation: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. COM (2013) 253, Brussels.

⁵² Marvin, S. , Luque-Ayala A. & McFarlane, C. (2016) Smart Urbanism: Utopian Vision or False Dawn? New York: Routledge

depends on the local environment. The roadmap must be perceived as a guide to help us be consistent and cover most aspects of our evaluation framework.

The roadmap incorporated the following phases that are linked with evaluation methods, and particularly with data collection methods, as you can also see in the figure below.



Indicative diagram of the continuity of data collection phases

In more details we have the following phases:

Usage data. Token related data, tags, interactions, gamification and comments will be collected from the CO3 platform throughout the whole implementation of the services along with relevant metadata. The analysis of data will help us spot significant aspects of the designed services, improve features during implementation and shape specific actions to investigate in the field. Quantitative analysis of the usage data will take place but we mostly expect to use them to provide insights in combination with the quantitative data in order to make

conclusions. You can find a comprehensive table of the Usage Data to be collected and analyzed in the Appendix.⁵³

- A. **Training Workshop.** Depending on the service, different workshops could be organised in each pilot to simulate the process of the proposed service to users and engage them. During these sessions, facilitators should shadow the interactions of participants without interfering and keep notes. These workshops could be combined with small focus groups or feedback sessions to get more specific impressions about the flow or other aspects of the service. In addition, these workshops can be part of the evaluation plan. Quantitative analysis, (Discourse/Thematic) will be implemented to analyze the data collected.
- B. **Preliminary Questionnaire.** These questionnaires should collect demographic data and quantitative data related to the conditions of the field in which the designed service will operate. It's also important to include qualitative open questions in order to get thoughts and impressions and of the participants about the project or things that could be designed differently. Our intention is to map in detail the present situation of the field, in order to identify the impact inflicted by the implementation of the service. The questionnaire will have both quantitative (likert-scale) and qualitative data and each section will be analyzed with the appropriate methods.
- C. **Shadowing on service implementation.** In this phase, we have to keep detailed notes about the progress of the service in the field and significant incidents related to the delivery of the service. It is important to write down possible changes on the behaviour of the users with the intention to intersect conclusions from usage data analysis with impressions and comments reported by facilitators in the field. To collect data we will shadow the service implementation without intervening. Another option is to ask simple questions to the users and collect data in scorecards. We will mostly work with a qualitative analysis of the reports here but there is also possible to work with qualitative data from the scorecard if the context requires so.
- D. **Final Questionnaire.** It is a tool also to collect quantitative and qualitative data. It will give us the opportunity to measure our variables and to designate impact and transformations on the ecosystem. We will develop sets of questions for services in the same category in order to have comparable data at the beginning and at the end of pilot implementation of the services. For this purpose it can work jointly with phase B.
- E. **Final Workshop.** This workshop will function as a more focused feedback session with users and people who participated in the phases of co-design and co-production of the service. These workshops will be organized upon specific points of interest that emerged during the implementation of the service. Quantitative analysis, (Discourse/Thematic) will be implemented to analyze the data collected.

⁵³ See Appendix B.Usage Data

F. **Experts interviews.** After the completion of the data collection and the categorization of our data, interviews with experts will help us to acknowledge our thematic lenses and give more depth to our analytical framework. It will also help us understand how specific stakeholders, especially Public Authorities, perceive the service and its sustainability. In this case we will work mostly with Discourse Analysis.

As we stated earlier this roadmap is indicative and its purpose is to cover the different qualities and dynamics which will emerge in cross pilot level. The phases of data collection might be rearranged to adapt to the needs of each pilot scenario.

Appendix

A. Pilot Services Reference

The CO3 Pilot Services and the sections where they are described in other deliverables.

Code Name	Title	D1.2	D3.1	D4.1
Athens.1	Groceries on Hold	5.3.4	3.1.1	5.4.1
Athens.2	Mapping community needs on empty buildings	5.3.5	3.1.2	5.4.2
Paris.1	Contributive Clinic	5.3.2	3.3.1	5.4.3
Paris.2	Digital urbanism in Augmented Schools	5.3.1	3.3.2	5.4.4
Paris.3	Blockchain Knowledge Registry	5.3.3	3.3.3	5.4.5
Turin	Augmented Commoning in CdQ	5.3.6	3.2	5.4.6

Table 1 - The CO3 Pilot Services

B. Usage Data

A comprehensive table with all the Usage Data to be collected and analyzed.

Code	Title	Description	Pilot Reference	System(s)
D.1	Age Group	under 18 / 19-29 / 30-39 / 40-49 / 50 +	Demographic	LiquidFeed back
D.2	Education Level	primary / secondary / bachelor / master or higher	Demographic	LiquidFeed back
D.3	Sex	Male / Female / Other / Rather not say	Demographic	LiquidFeed back
D.4	Area	open question / rather not say	Demographic	LiquidFeed back
W.1	Volume of tokens raised	Volume of tokens raised in the pilot or in specific Wallets/User Group	A.1.1, P.3.1, T.2.1	Wallet
W.2	Volume of tokens consumed	Volume of tokens consumed in the pilot or by specific Wallets/User Group	A.1.3, T.2.2	Wallet
W.3	Time of tokens raised	The timeframe of tokens raised in the pilot or by specific Wallets/User Group	P.3.2	Wallet
W.4	Average purchase of tokens	The average purchase of tokens by User Groups	T.2.2	Wallet
W.5	Average balance	The average balance of wallets in the pilot or in specific Wallets/User Group	T.2.2	Wallet
W.6	Average time to spend	The average time in which the tokens are spent in the pilot or by specific Wallets/User Group	T.2.2	Wallet
W.7	Number of users that use tokens	Number of users that use tokens and distinction by user group	T.2.6 , T2.12	Wallet
W.8	Number of user in more than on ACA	Number of users that use tokens in more than one ACA	T.2.7	Wallet
W.9	Tokens consumed or exchanged for money	Usage of tokens by wallets (redemption or exchange)	T2.11	Wallet
W.10	Transaction in User Group	Number of transactions within User Groups	T.2.13	Wallet
W.11	Users of specific token type	Number of users of specific token type or combination of types	T.4.1 , T.4.3.	Wallet
W.12	Volume height of specific token type	How many tokens have been created for a specific token type	T.4.3	Wallet
W.13	Issuers that accept 3rd party tokens	Number of Issuers that accept tokens other than their own	T.2.14	Wallet
FL.1	Number of AR	Number of items mapped through the	P.2.1, P.1.2,	AR App ,



	objects placed	AR App in FL	T.1.2	FirstLife
FL.2	Items mapped	The content of the items mapped through the AR App in FL	A.2.1 , P.1.5, P.1.6, P.2.2	FirstLife
FL.3	Items mapped by type	They type of items mapped/placed in AR	A.2.1, T.3.3	AR App , FirstLife
FL.4	User views by item	Views of single users per item	A.2.1, T.1.3, T.3.2.	FirstLife
AR.1	Logged Time	Time in the platform linked with items placed during that time	P.1.1, P.2.3	AR App , FirstLife
AR.2	Number of ACA created		P.2.4	AR App
AR.3	Views and Shares of AR Content		T.1.1	AR App
AR.4	Number of POI (point of interest)	Number of points in the map where AR objects have been placed	P.2	AR App , FirstLife
AR.5	Number of 3D models created for ACA	Number of different 3D models loaded in the app per ACA	P.2	AR App , FirstLife
LF.1	Number of Proposal and comment	Numerical analytics on the discussions made in LF for the items in FL	A.2.3, P.2	LiquidFeed back

Table 2 - Usage Data

