

A Framework for Evaluating the Implementers' Experience in Making Existing Products Accessible: The Prosperity4all Approach

Katerina Toulou¹(✉), Maria Gemou¹, Till Riedel², Maria Panou¹,
and Evangelos Bekiaris¹

¹ Centre for Research and Technology/Hellenic Institute of Transport
(CERTH/HIT), Athens, Greece

{toulouk, mgemou, mpanou, abek}@certh.gr

² Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
riedel@teco.edu

Abstract. Prosperity4All is a continuous and dynamic paradigm shift towards an e-inclusion framework building on the architectural and technical foundations of other Global Public Inclusive Infrastructure (GPII) projects aiming to create a self-sustainable and growing ecosystem where developers, implementers, consumers, prosumers and other directly and indirectly actors (e.g. teachers, carers, clinicians) may play a role in its viability and diversity. An agile and dynamic approach is adopted in three evaluation phases, starting with formative evaluations with five internal implementers leading to more summative techniques towards the final evaluation phase where more ($n = 25$) and external professionals will use the tools and resources available in the project's repository (DeveloperSpace) to improve and enhance their own products and services. The evaluation approach for the implementers considers three dimensions: (a) the project's Key Performance Indicators (KPIs), (b) technical validation activities prior evaluation, and (c) three evaluation phases followed by a final impact assessment.

Keywords: Inclusive design · Implementers · Evaluation · Accessibility · Ecosystem

1 Introduction

Prosperity4All is a continuous and dynamic paradigm shift towards an e-inclusion framework building on the architectural and technical foundations of other Global Public Inclusive Infrastructure (GPII) projects by creating a self-sustainable and growing ecosystem, where developers, implementers, consumers, prosumers and other directly and indirectly actors (e.g. teachers, carers, clinicians) may interact with and play a role in its viability and diversity. The Global Public Inclusive Infrastructure (GPII) is a project of Raising the Floor, a consortium of academic, industry, and non-governmental organizations and individuals (<http://gpii.net/>). The GPII will

combine cloud computing, web, and platform services to make access simpler, more inclusive, available everywhere, and more affordable. When completed it will provide the infrastructure needed to make it possible for companies, organizations, and society to put the web within reach of all - by making it easier and less expensive for consumers with disabilities, ICT and AT companies, Public Access Points, employers, educators, government agencies and others to create, disseminate, and support accessibility across technologies.

In particular, the aim of Prosperity4All is to provide an infrastructure for the development of an ecosystem by employing modern and new techniques, like crowdsourcing and gamification, to enable new strategies for developing accessibility services and introduce a new approach to accessibility solution development. This ecosystem will allow seamless, efficient, cost-effective and unobtrusive communication between developers, implementers, consumers and prosumers. Consumers will be able to communicate with developers and implementers for ordering personalized and customized products and solutions (e.g. web-based business solutions customized for visual impaired users). However, with such diversity comes complexity that substantially affects the designing and planning process for the respective evaluation approach and framework.

The evaluation of the ecosystem will be achieved through impact estimations of its deployment. Before reaching the point to estimate small or large potential impacts, actual evaluations will be carried out in three pilot sites in Europe; Austria, Germany, Greece, Spain with real users and implementers. The evaluations with implementers precede any testing with real end-users. The final evaluation phase with implementers will be the first evaluation phase with end-users. Evaluations with implementers will be performed with at least thirty users in different sites, including both internal to the project participants and externals for the second and third evaluation phase. The objectives and the Key Performance Indicators (KPIs) were the driving forces for drafting the evaluation questions to be considered and set.

1.1 The Overarching Evaluation Questions

Defining the questions to ask was a first step. The evaluation questions had to accommodate for the project's KPIs and the latter reflected the objectives. All three were mapped before overarching questions were prepared. Thus, a top-down approach was followed for the evaluation questions of the framework.

1. Are the tools/resources for Developers (DeveloperSpace and all of the frameworks, components, marketing tools, etc.) usable by and useful to developers/implementers (both internal developers and external developers, implementers)?
2. Do the tools/resources help implementers in their work or decrease cost to develop or increase market size/share? OR increase profits?

Evidently, the evaluation focusses not only on the utility and use of these tools and resources but also on their cost-efficiency in their everyday work. Therefore, implementer's previous experience with and involvement in accessibility work and projects is of importance. A bottom-up process is applied for preparing the actual evaluation materials where specific instruments are selected.

1.2 The Evaluation Framework for Implementers

Developing an inclusive and human-oriented framework that will adapt dynamic and agile methods and will be embedded in the development lifecycle, to the extent this is possible, is a challenging endeavor for planning the evaluation but as well as collecting data and draw inferences on the outcomes. Early evaluations are mostly formative leading later on to more summative efforts. The final evaluation phase will coincide with the first iteration with end-users and then their inter-play will be captured with pluralistic techniques.

There is keen interest in identified how cost-efficient and viable will be this ecosystem for professionals working in diverse areas (e.g. web developers, hobbyists, etc.) in order to offer customized and personalized solutions to people with diverse and sometimes complex accessibility needs, aiming to address the tails-of-the-tails of populations of users they might be isolated by existing practices and offered market-place solutions. The evaluation framework addresses different users-actors based on their functional role which they may play in the ecosystem. The roles might as well be interchangeable when the ecosystem will be deployed, after all evaluations finish and optimization is achieved; wherever relevant and applicable.

Building up the evaluation framework requires knowledge in the areas of traditional usability and user experience testing and insight in customer perception and e-commerce marketing analytics. Most of the project's applications and services are already offered to consumers and therefore they will not be evaluated per se. On the contrary, the tools developed or improved during the project - tools and frameworks with Graphical User Interfaces for Development (IDEs), building blocks and frameworks (with no graphical interfaces) for developers (APIs), and web-based developer resources- will be available at a specially design repository, the DeveloperSpace.

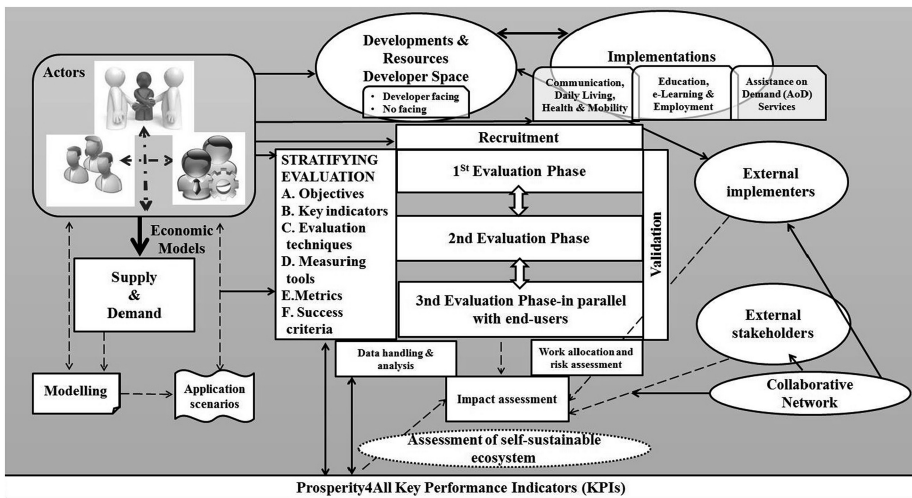


Fig. 1. The implementers' evaluation framework

The three evaluation activities will be harmonized with the enhancing development cycle. Firstly, the implementers will search the DeveloperSpace to find the appropriate tools and resources in order to add functionalities to their products and services for making them more accessible or adding functionalities that will make them even more accessible to users with other accessibility needs. The evaluation framework is “surrounded” by other activities (Fig. 1) as the inter-connections and inter-dependencies between the evaluations and the following project activities are necessary:

- (a) the ecosystem’s business cases defined by the demand and supply chains for the future actors in the alive ecosystem;
- (b) the tools (e.g. developer-facing with interfaces like certain APIs) and resources chosen and used by implementers that reside in the DeveloperSpace; and
- (c) the actual products that will be improved, representing different areas of interest and life activities (e.g. business, health, education).

The logical model prepared for the evaluation activities with implementers includes the *input* namely the actors and the developments, the *process* being the evaluation activities including recruitment and technical validation whenever and wherever relevant, and the *outcomes* being the collection of the indicators by using formative methods in the first iteration (i.e. pluralistic walkthroughs and workshops with emphasis on implementers’ decision-making processes and utility of tools) and more summative in later stages (cost and time efficiency estimations and matching the expectations of developers with different levels of experience in accessibility with their post-responses).

2 Methodology

The methodological approach adopted includes three inter-dependent dimensions: (a) objectives as set by the Key Performance Indicators (KPIs) of the project, (b) technical validation prior any testing takes place, (c) three evaluation phases, and a final impact assessment. The evaluation framework led to the design of a logical model specifically for testing with developers taking into consideration these three evaluation dimensions and two meta-evaluation aspects which are important for the self-sustainability of the ecosystem and, thus, for its prosperity; an agile feedback loop utilizing contemporary web tools such as JIRAs and a meta-evaluation assessment carried out after the end of each phase (i.e. structured lesson learnt method based on pre-defined mitigation planning). The implementers’ logical model addresses the following for each category of developments (IDEs, APIs, etc.): (a) **a higher objective**; e.g. matching of notations and graphical elements regarding relevant user, (b) **indicators/constructs**; e.g. 12 cognitive dimensions; abstraction gradient, closeness of mapping, consistency, etc., based on Cognitive Dimensions Theory [1], (c) **evaluation technique** (e.g. scenario-based cognitive analysis); (d) **evaluation instrument/tools**; e.g. cognitive dimensions’ questionnaire [2], and **success thresholds & criteria**; e.g. approximate matching dimensions.

The developers are interested in how these tools match the process and applications (matchmaking) and how much they will save in money and time when using those

tools (cost-effectiveness). These are also considered, as we want to understand how the developers will reach the decision to use the tools (i.e. revealing the decision making process) and which is the reference case (i.e. their professional preferences and decisions to use certain tools over others in their work environment). The evaluation focuses on measuring how usable and useful will be the DeveloperSpace and all of the frameworks, components, and tools for internal and external implementers, with consideration for how helpful they will be, how much they will decrease the cost to develop, and how much they will increase profits and market size. The last two aspects will be addressed by the impact assessment carried out after the deployment of Prosperity4All ecosystem. The professional experience of developers and implementers participating in the pilots in inclusive design and accessibility is taken into consideration. Three iteration phases will be carried out starting with a small group of internal implementers (N = 5) for the first iteration phase and gradually including external implementers in the last two (N = 25). Peer reviews and automatic documentation improvements are common qualitative methods for evaluating solution for developers. These methods have restricted transferability and therefore potential validity. Instead, heuristic walkthroughs and relevant formative techniques will be used in the first iteration. Summative evaluation will mostly take place in the iterations to follow; especially when the Prosperity4All platform will be deployed and analytics will be gathered from real life interaction with it.

2.1 Actors

Implementers are the Prosperity4All actors who will incorporate the tools and resources offered to them in the DeveloperSpace in order to make their applications and services more accessible or to improve the user experience of already accessible applications. Implementers are both the internal implementers and external professionals who might be freelancers or even companies, service providers, and other groups as identified in the list of actors (Table 1).

In these user groups, developers who will directly add outcomes to their applications and use existing resources during the improvement process (enhancement

Table 1. Primary categories of implementers

<p>1. AT Developers</p> <ul style="list-style-type: none"> • Hardware AT • Software AT (install) • Web/Cloud AT <p>2. Mainstream Developers</p> <ul style="list-style-type: none"> • Desktop Applications • Web/Cloud Applications • Cloud Service Providers • Hardware/Appliances • Mobile Applications <p>3. System Integrators</p>	<p>4. Researchers</p> <ul style="list-style-type: none"> • University/College/Tech-Institute • Other technology-oriented education <p>5. Community Developers</p> <ul style="list-style-type: none"> • Micro Service Developers • Friends and Family • Prosumers <p>6. Service Delivery Professionals</p> <ul style="list-style-type: none"> • Clinicians • Teachers
--	--

development lifecycle), their implementation will also evaluate this work and implementers will evaluate the utility –among other attributes–of the resources they will choose to “accessibilize” their products. The participants are sought to be representative of a bigger group of “implanters” that need to be considered in the wider scope (Table 1).

Other stakeholders may be influenced by the implementer’s perspective of the DeveloperSpace; this particularly includes the government. Governmental agencies are setting the regulatory frame for many implementations as well as do procurement officers or decision makers. They highly influence what will be considered for implementation. The evaluation acknowledges that roles may be fluid, so that also consumers with decision making powers have influence on implementation decisions and that particularly prosumers are interesting stakeholders in the realm of accessibility. One underlying assumption of the evaluation framework is, however, that the implementers’ perspective on Prosperity4All is common but very heterogeneous for all the stakeholders.

Examples of Generic Personas and Application Scenarios for Implementers.

Based on the list of actors, application scenarios like short stories are created for three groups of actors which belong to the producer category (i.e. people who produce products, create applications, improve services, etc.) and have a direct impact to the Prosperity4All developments (i.e. belong to the family of Producer of Things (PoTs) scenarios). They were created for three different value propositions (i.e. reasons to join platform). At this stage, scenarios are characterized by the functional role of the stakeholder, the value proposition (broad), and the family of scenarios it belongs.

The value of these scenarios for the evaluation framework lies in the fact that they provide insight in the many types of actors that they could be involved as implementers, the way they can work and collaborate, and the variations in their expertise, knowledge and even the areas of interest within accessibility. They have a rather illustrative and communicative value between economic modelling and evaluation than a direct application and implementation to any measurable conditions and aspects.

The personas include functional elements (e.g. what the identified persona is doing with the system) accompanied by a short application scenario for testing purposes.

Three potential generic personas and application scenarios are based on initial ideas of the how main actors will interact with the system-in still a fragmented style- but focusing mainly on the story about who the user of a particular technology is, what they want, what they know.

Persona 1: Actor – Producer-Economics: GUI adaptation of route guidance system for visually impaired users (Support independent living).

Simon is a developer (Actor -Producer – supply-end of chain) who has long been working in making accessible applications for many years. He has worked in a large company for many years and lately he is interested in navigation support systems for marginalized user groups such as people with visual impairments. He found out about the Prosperity4All multi-sided platform via blog for developers he often visits and receives the Newsletter. When he visits the developer part of the platform he is unsure about which component of the DeveloperSpace is more appropriate for what is looking

for to do. He checks the link and visits the Prosperity4All training platform. He selects the curriculum for external implementers and specifically the course on adapting GUIs for visually impaired users especially for navigation support software.

He then selects the component for changing the interface of the routing guidance system and makes it available to the platform for users to buy. There is also an option for the user to ask for a specific customization to be made and there is an opportunity to hold a discussion with the developer prior the purchase.

Application of scenario for testing: The implementer will adapt the GUI interface of the route guidance system for visually impaired users. Testing at early stages of development will be performed together with other low or medium fidelity prototypes.

Persona 2: Actor – Producer-Law: Making accessible learning materials (Support independent education and work).

Carla (Actor – Producer – supply end of chain) is a freelancer who is currently collaborating with a large public library aiming to make their digital resources accessible to blind and visually impaired users. She visits the Prosperity4All platform and accesses the part for developers and implementers in order to find relevant resources for her work. The training videos were very helpful and she found numerous resources about different screen readers and their implementation to the vast and diverse digital books and information available. The workload is huge but still the resources and tools available at the Prosperity4All platform will assist Carla by saving-time looking for methods and tools in the internet and increasing her potential and knowledge in the accessibility domain.

Application of scenario for testing: The implementer will select a tool to enhance the accessibility of digital documents to be accessible by blind and visually impaired users (e.g. either one or two screen readers). Testing at early stages of development will be performed together with other low or medium fidelity prototypes.

Persona 3: Actor – Producer-Ethics: Adaptation of Assistance on Demand (AoD) services for older people (Support inclusion of lower or no literacy computer users).

Nick (Actor-Producer-supply end of chain) is working as a developer and IT specialist in a national bank branch. He is also a volunteer at the regional Elderly Centre near his home. He is deeply concerned about older people and their lower digital literacy. He is helping them to learn how to use computers. He wants to find a way to help older visitors use the website of the elderly center. He is teaming up with a friend who is actually working as social worker at the center and is pretty aware of the problems older computer users might face and he is just an enthusiast (i.e. he is an amateur software designer). A friend informed him about the Prosperity4All platform and the availability of the AOD framework for enhancing the existing AOD set up of the service provided by the Elderly Centre in order to provide appropriate and adequate technical support to lower digital literacy older users. Their work aims to increase independent use of computers by the users.

Application of scenario for testing: The implementer will use the AoD infrastructure to enhance the AoD services and make them more accessible. Testing at early stages of development will be with other low or medium fidelity prototypes.

2.2 Tools, Resources and Products

The DeveloperSpace repository includes the tools and any relevant additional documentation (e.g. instructions, manuals, etc.). At the very early stages of the project, many of these tools and resources are available as prototypes, mock ups or even as a proof-of-concept. While most of the applications and services have been evaluated already within the scope of other projects, enhancing them is “another story to be told”. These tools will be used to make accessible or improve the existing accessibility of different products and services, covering needs from many areas of daily activities (such as communication, education, health, and employment).

The tools and resources used by implementers fall in to the following three main categories and will be:

- Web-based Developer Resources and Assistance on Demand (AoD) services.
- Tools and Frameworks with Graphical User Interfaces for Development (IDEs).
- Building Blocks and Frameworks (with no graphical interfaces) for developers (API).

Those three categories are driven by both the categories of different outcomes (components, tools, services, and infrastructure) and practical considerations and needs for testing and evaluation. Both outcomes and implementations can belong to multiple categories. The implementers will use tools from these categories to improve, enhance and add new functionalities to more than ten existing products falling into the following three main types:

- Communication, Daily Living, Health, and Accessible Mobility.
- Education, eLearning, Business and Employment.
- Assistance on Demand (AoD) Services.

One important consideration was made within the evaluation framework and that was to evaluate the specific effect of Prosperity4Aall as much as possible. While most of the applications and services were evaluated already in the scope of other projects, enhancing them is “another story to be told”. It is very important that human factor evaluation also focusses on the unique prosperity propositions that come through the project and the exposure of the tools within an ecosystem. Therefore it is important to understand that all interactions between developers and implementers are made through exactly that evolving ecosystem. Evaluation will be carried out for both the DeveloperSpace (where all tools will be available) and the use tools, resources and applications (internal and external) in the context of the DeveloperSpace.

Particularly for the infrastructure of the DeveloperSpace, the project will develop multiple developer-facing components that are exposed to implementers (there is also user facing components exposed to end-users that are not part of the testing methods). Many of those outcomes will be presented as web-based developer resources. Most prominent example is the component listing (repository) that will be a directly visible outcome. In those cases, proven user experience methodology can be applied. The user model of a developer is different from an end-user in the domain, however, particularly here transition between roles need to be considered for certain stakeholder classes (for user-programmers).

Particularly for the first and second category of tools, the matchmaking aspect is becoming of further importance. While many of the web-based resources will be an entry-point for many types of stakeholders, the picture differentiates quickly after that. Particularly there will be “no-one-fits-all” usability for components. The goal of Prosperity4All is to enable the selection of fitting components and furthermore the fitness of the components for relevant stakeholders, which differ for component to component.

Because often it is not easy to get a summative picture regarding components, services and tools and implementations there will be a two stage process. The first stage will be a match-making inside the project via the DeveloperSpace. This matchmaking already takes the usability of the web resources into account. After this initial match-making, the hypothesis is built that the selected implementation should be fitted to the implementer that selected it. In the second step we are particularly evaluating the usability based on this assumption and use the evaluation also as a formative tool to improve the tools to become better. For those evaluations, established human factor evaluation techniques can be partially applied; particularly if a tool exposes a graphical interface to the developer.

The following high level objectives will be measured:

- Strengths and weakness of used tools and resources.
- Cost-efficiency perceived measures (e.g. considering time and effort compared to current practices; consideration for experience in accessibility is important)
- Improved user experience.
- Utility, usefulness and learnability.
- Acceptance compared to current practice.
- Matching of notation and graphical elements regarding relevant user and development activities.
- Freeing of developer to concentrate on creative aspects of the process.
- Global developer experience with focus on perceived attractiveness (for developer-facing tools and efficiency of offered and available tools and resources).
- Matching of offered tools and resources to needs and requirements.
- Decision making process and changes to it.
- Knowledge and experience driven attitude and productivity.
- Willingness to use and apply in future.

They are high categories which are stratified further into simpler constructs and will be matched to further methods and indicators during the lifetime of the project.

2.3 Conditions and Techniques

Testing with implementers will be primarily carried out in three contexts: (a) in their own work environment with real use of components and tools (group assessment, i.e. peer heuristics), (b) remote testing (remote data gathering), and (c) face-to-face qualitative assessment. Considerable part of testing will be carried out in their own

environment gathering mostly qualitative data. Focus groups will be carried out with implementers in small groups (5–8 participants) which could be organized in parallel with demo workshops. The focus groups will provide enriching data to the interviews held with participants and the rest of collected data. This is relevant to data triangulation and filling the “gaps” of other methods of data acquisition.

Conditions will vary between phases and among users. Each user might implement different outcomes for making an application or service accessible. During the first phase, implementation will be in some cases emulated as non-functional versions of tools might only be available (e.g. mock ups, paper prototypes). Another sample of implementers should be anticipated for the impact assessment. They will probably remotely assess the Prosperity4All ecosystem as part of impact assessment. These users will freely interact with the platform and evaluation will involve a real-life assessment of the ecosystem.

The conditions of testing are based on application scenarios which will serve the requirements of the evaluation framework.

Apart from traditional testing aspects, there are therefore two key features considered in the evaluation framework: (a) *matchmaking* – how tools fit the process and applications, (b) *cost-effectiveness* - of the use of specific tools in order to reveal how implementers are driven to choice of tool(s) (i.e. elaboration on decision process in relation to reference case). For the later multi-criteria analysis will be used for certain number of implementers. Tests with the applications and services will investigate the applicability and usefulness of the technology infrastructure. Most of the implementations are applications and services already offered to consumers and therefore they will not be evaluated per se. On the contrary, many of the tools and the DeveloperSpace will be developed within the project and therefore many of them will be offered as prototypes, mock ups or even proof-of-concept at the very early stages of the project.

3 Conclusion

The user model of a developer is different from an end-user in the domain, however, particularly in this case transition between roles needs to be considered for certain stakeholder classes (for user-programmers). This work will provide an understanding of implementer’s experience. Developers and implementers usually are on the receiving end of the evaluation process (i.e. receiving feedback) but in this case they will be actively communicating their experience with certain tools and products.

In a world of increasing ICT expertise, and greater future overlap among professional disciplines, the boundaries between users and developers are expected to become difficult to draw and balanced knowledge of both ends is of core importance for valid and reliable inferences.

Acknowledgments. This paper presents work carried out in Prosperity4All project. This project has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 610510. <http://www.prosperity4all.eu/>

References

1. Blackwell, A.F., Green, T.R.G.: A cognitive dimensions questionnaire optimised for users. In: Proceedings of the Twelfth Annual Meeting of the Psychology of Programming Interest Group (2000)
2. Green, T.R.G.: Cognitive dimensions of notations. In: Sutcliffe, A., Macaulay, L. (eds.) *People and Computers V*, pp. 443–460. Cambridge University Press, Cambridge (1989)