Living Labs and Open Toolkit for the co-design of accessible IT systems and tools

PROJECT
Living Labs and Open Toolkit for the co-design of accessible IT systems and tools
AGREEMENT NUMBER
LC-01624438/101018048

D.5.2.2 – Final Evaluation Report

SUBMISSION DUE DATE	DELIVERABLE VERSION
30/6/22	2

Co-funded by the European Union

DELIVERABLE TITLE	Final Evaluation Report
DELIVERABLE No.	D.5.2.2
Deliverable Version	2
Deliverable Filename	Final Evaluation Report
Nature Of Deliverable	Report
Number Of Pages	74
Work Package	WP5
Partner Responsible	Arcola Research
Author(s)	Joe Cullen
Contributor(s)	All partners
Reviewed by	Ioannis Poultourtzidis
Approved by	Panos Bamidis, Project Coordinator

PROJECT FULL TITLE	Living Labs and Open Toolkit for the co-design of accessible IT systems and tools
Start Of Project	1 April 2021
Duration	12 months
Project URL	**https://liveit-project.eu/**
Disclaimer	The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Table of Contents

GLOSSARY AND LIST OF ACRONYMS **4**

REPORT SUMMARY **7**

1 INTRODUCTION **8**

1.1 PURPOSE AND SCOPE 8
1.2 INTENDED AUDIENCE 9
1.3 STRUCTURE OF THE DELIVERABLE 9
1.4 RELATION TO OTHER WPS & TASKS 9

2 EVALUATION APPROACH AND METHODOLOGY **9**

2.1 THE 'OBJECT' OF EVALUATION 9
2.2 EVALUATION IMPLEMENTATION IN LIVE IT 11
2.3 EVALUATION METHODOLOGY 12

3 EVALUATION RESULTS **16**

3.1 OUTPUTS ANALYSIS 16
3.2 EVALUATION OF THE CO-LABS 17
3.2.1 Mapping and assessing digital challenges *17*
3.2.2 Co-Lab activities implemented *19*
3.2.3 Characteristics of Co-Lab participants *20*
3.2.4 Co-creating solutions to digital challenges *22*
3.2.5 Tools evaluation *25*
3.3 EVALUATION OF THE MAKERSPACE AND HACKATHONS 38
3.4 EVALUATION OF THE OPEN TOOLKIT 44
3.5 ECONOMIC ANALYSIS 54

4 THE BIG PICTURE: LIVE IT OVERALL SUMMATIVE EVALUATION **58**

4.1 THEORY OF CHANGE ANALYSIS 58
4.2 KEY EVALUATION MESSAGES AND IMPLICATIONS 63

5 REFERENCES **74**

GLOSSARY AND LIST OF ACRONYMS

Term	Definition and source
Action research	Practice based research, which seeks to end the dislocation of research from practice and enhance the position of research as a direct mechanism for change and improvement. Link
Attribution	The ascription of a causal link between observed (or expected to be observed) changes and a specific intervention. Note: Attribution refers to that which is to be credited for the observed changes or results achieved. It represents the extent to which observed effects can be attributed to a specific intervention or to the performance of one or more partner taking account of other interventions, (anticipated or unanticipated) confounding factors, or external shocks. Link
Behavioural additionality	Changes in beneficiaries' behaviours resulting from an intervention. Link
Contribution analysis	Contribution Analysis is an approach for assessing causal questions and inferring causality. It offers a step-by-step approach designed to help managers, researchers, and policymakers arrive at conclusions about the contribution their program has made (or is currently making) to particular outcomes. The essential value of contribution analysis is that it offers an approach designed to reduce uncertainty about the contribution the intervention is making to the observed results through an increased understanding of why the observed results have occurred (or not!) and the roles played by the intervention and other internal and external factors. Link
Counterfactual	The situation which would have arisen had the intervention not taken place. Link
Critical Success Factors (CSF's)	'The critical areas whose high performance or success is important' and also 'the steps taken to succeed' (Rockart, 1979)
Ex ante evaluation	An evaluation conducted before the implementation of an intervention.
Indicator	A characteristic or attribute which can be measured to assess an intervention in terms of its outputs or results. Output indicators are normally straightforward. Result indicators may be more difficult to derive, and it is often appropriate to rely on indirect indicators as proxies. Indicators can be either quantitative or qualitative. Context indicators relate to the environment for a project or programme. Link
Formative evaluation	Evaluation which is intended to support programme actors, i.e., managers and direct protagonists, in order to help them improve their decisions and activities. It mainly applies to public interventions during their implementation

	(ongoing, mid-term or intermediate evaluation). It focuses essentially on implementation procedures and their effectiveness and relevance. **Link**
Key Performance Indicators	Key Performance Indicators (KPI's) make the connection between the CSF's and the KRIs. They track the actions between the CSF's and the KRIs. So, first, they have to measure a process. Second, they have to be key - i.e., the only measures that are essential to demonstrate progress towards 'results'. Third, they have to measure 'live' data - i.e., the information source used to measure process and progress is continually generating updated information. Fourth, they need to reflect 'context'. Fifth, they have to be 'metrics' - i.e., a quantifiable measure that can demonstrate progress either from a baseline or in context.
Key Results Indicators	Measure the 'results' (effects) of steps taken to succeed that are carried out in terms of the 'end result'
Participatory evaluation	Evaluative approach that encourages the active participation of beneficiaries and other stakeholders in an evaluation. They may participate in the design and agenda setting of an evaluation, conduct self-evaluations, help gather data or help interpret results. **Link**
Process evaluation	Focuses on learning about, and potentially improving, delivery. Tavistock Institute
Summative (or ex post) evaluation	It is conducted after completion and for the benefit of some external audience or decision-maker (e.g., funding agency, historian, or future possible users). (Scriven M., Evaluation Thesaurus). **Link**
Theory based evaluation	Theory-based approaches to impact evaluation allow for a systematic articulation and testing of the assumed connection (i.e., the theory) between an intervention and the anticipated impacts. The focus of theory-based evaluations is not only on understanding whether an intervention has worked but on why and under what conditions change has been observed. Tavistock Institute
Theory of change	Theory of Change is a systematic and cumulative study of the links between activities, outcomes, and context of an intervention. It involves the specification of an explicit theory of how and why an intervention might cause an effect which is used to guide the evaluation. It does this by investigating the causal relationships between context-input-output-outcomes-impact in order to understand the combination of factors that has led to the intended or unintended outcomes and impacts. Theory of Change therefore tests, and normally develops the implementation theory of an intervention and allows this to be modified or refined through the evaluation process. Tavistock Institute

Disclaimer Statement

The content of the publication herein is the sole responsibility of the publishers and does not necessarily represent the views expressed by the European Commission or its services. While the information contained in the document is believed to be accurate, the authors(s) or any other participant in the LIVE IT consortium make no warranty of any kind with regard to this material including, but not limited to the implied warranties of merchantability and fitness for a particular purpose. Neither the LIVE IT Consortium nor any of its members, their officers, employees or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein, or direct, indirect, special, incidental or consequential damages in connection with the furnishing, performance, or use of this material.

Statement of Originality

This Deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Copyright

© The LIVE IT Consortium, consisting of:

- Aristotelio Panepistimio Thessalonikis (AUTH), Greece
- Arcola Research LLP (ARC), United Kingdom
- National University Of Ireland Galway (NUIG)
- Universidade Catolica Portuguesa (UCP), Portugal

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License, 2017. For details, see **http://creativecommons.org/licenses/by-sa/4.0/.** The text, figures and tables in this report can be reused under the provisions of this License. Logos and other trademarks are not covered by this license. The content of the documents marked as restricted or confidential are not to be disclosed externally without prior written consent from the LIVE IT Consortium.

Acknowledgements

The LIVE IT team would like to thank all of our collaborating partners, especially those organisations working with us on the ground in developing the Co-Labs, and all of the LIVE IT 'users' – people with cognitive disabilities – for giving their time and insights to contribute to this research.

REPORT SUMMARY

This document presents the final evaluation of the LIVE IT project. It speaks mainly to the 'accountability' and 'learning' purposes of evaluation. It is a 'summative' evaluation that seeks to provide evidence on whether the project's goals have been met and how. It presents the overall approach to evaluation and the methodology used to evaluate LIVE IT, the key results of the evaluation and an overview of the main evaluation conclusions and implications.

The overall approach applied in the LIVE IT evaluation is based on an adaptation of the 'realist evaluation' approach. A realist approach is theory-based and is essentially about testing a theory about what 'might cause change', even though that theory may not be explicit. LIVE IT's Theory of Change maps the project 'change journey' so we can see the connections between the 'presenting problem' it wants to solve, the expected impact on that problem at project end and everything that's supposed to happen in between.

- Analysis of the project outputs shows that, across most evaluation measures, LIVE IT has exceeded its targets, in particular for baseline research, the number of participants working in the LIVE IT 'Co-Labs', the accessibility tools evaluated in Co-Lab experiments and the outputs produced in the Labs.
- LIVE IT's outreach and dissemination activities present a mixed picture. On the one hand, the project reached or exceeded its targets in terms of webinars and local seminars delivered, conference presentations and peer-reviewed journal articles produced and attendance at the final conference. However, LIVE IT had limited success in reaching its broader target audience of stakeholder communities. In particular the level of involvement in the Makerspace was very low.
- The preliminary work carried out through baseline research, followed by intensive work with participants in the Co-Labs led to the mapping and analysis of around 85 challenges that prevent people with cognitive disabilities fully benefiting from digital life. These challenges were mapped against 1200 tools with the potential to deliver solutions to these challenges.
- Subsequent co-creation work carried out in the Co-Labs, involving 355 people with cognitive disabilities as co-designers and co-developers exploring the efficacy of a range of these tools, covered some 351 experiments in 72 different digital challenge scenarios. 125 tools to improve web accessibility for people with cognitive disabilities were systematically evaluated in the Co-Labs, on the basis of usability and usefulness and 'emotional response'. The results of this work fed into the development of the LIVE IT 'Open Toolkit'.
- In parallel, five Hackathons were implemented over the project lifecycle. 200 people participated in the Hackathons, in 38 different teams. These teams carried out further experiments with the suite of assistive tools, and with the LIVE IT 'Open Toolkit'. As a result of the Co-Lab and Hackathon work and, to a lesser extent the work done in the LIVE IT Makerspace, around 200 'interventions' were co-produced involving

recommendations for changes to assistive tools, ideas for new tools and improvements to the Open Toolkit.

- The Open Toolkit is a major resource for a broad range of stakeholders, including people with cognitive disabilities, the organisations that represent them, the IT industry and policymakers. It distils and valorises all of the evidence and learning from the project and provides a range of tools to support web accessibility, including an 'Advisor Tool' to assist users in making evidence-based choices on which assistive tools to use. Evaluation of the Open Toolkit was very positive.
- An economic evaluation of LIVE IT, using a 'cost consequence analysis' (CAA) methodology suggests that the potential economic benefits associated with increasing web accessibility for people with cognitive disabilities are considerable, going forward.
- Theory of change analysis shows that LIVE IT has progressed to the 'behaviour change' point of a 'change journey' that, ultimately, could lead to systemic, disruptive and transformational change in the way digital tools and services are designed. This position is what might be expected for a preparatory and pilot action like LIVE IT. The evaluation shows that the project has provided a justification for further actions in this field; developed a significant evidence base on what works; developed extensive resources to help future researchers, policy makers, IT developers and service providers change the way tools and services are designed and laid the foundations for the dissemination and application of the learning derived from its activities.

1 INTRODUCTION

1.1 PURPOSE AND SCOPE

This document presents the final evaluation of the LIVE IT project. Although not specified in the project Description of the Action (DoA) as a formal deliverable, the DoA included as 'evidence of success' an 'Evaluation Report year end specifying key outcomes for participants' (p. 3). This evaluation report is therefore provided as a follow up to D5.2 – Interim Progress Report, as part of the activities carried out in work package 5 – 'Dissemination and Management'.
Evaluation in LIVE IT had four main purposes:

- A developmental purpose – contributing to the project design and its outputs. For example, we developed a methodology to assess the robustness and quality – evidence effectiveness - of the examples of good practice cases identified in the 'state of the art review' in work package 1, and which ultimately were included as a 'Catalogue' of good practice cases in the 'Open Toolkit'.
- An operational purpose – contributing to monitoring the progress of the project and highlighting issues and challenges that may need to be addressed over the project life cycle. This was mainly implemented through a 'process dashboard' that tracked progress against key project outputs – for example number of participants engaged in

Co-Lab activities – and key performance indicators (KPIs) – for example change in number of experiments carried out in the Co-Labs against baseline.

- An accountability purpose – collecting evidence on the results of the project, particularly its outcomes and potential impacts, and on whether the project has delivered value for money.
- A learning purpose – including supporting reflection on the results of the project as they develop, through partner review workshops, delivering a programme of webinars and other events, and applying the evaluation results to promote the sustainability of LIVE IT beyond the formal project end.

This Final Evaluation Report speaks mainly to the 'accountability' and 'learning' evaluation purposes. It is a 'summative' evaluation that seeks to provide evidence on whether the project's goals have been met; the way the LIVE IT project was implemented; the quality of its outputs and outcomes; its main achievements together with likely long(er) term impact; the extent to which LIVE IT has contributed to the evidence and knowledge base relating to digital inclusion, in a way that contributes to improving policy and practice, especially as regards people with cognitive disabilities; the sustainability of the project.

1.2 INTENDED AUDIENCE

This document should be used as a reference by:

- all Consortium Partners of LIVE IT
- the European Commission and external reviewers
- LIVE IT stakeholders, including policymakers, people with cognitive disabilities and organisations representing them, the research community, IT designers and developers, digital service providers.

1.3 STRUCTURE OF THE DELIVERABLE

Following this introduction, Section 2 presents the overall approach to evaluation and the methodology used to evaluate LIVE IT. In Section 3 we present the key results of the evaluation. The concluding section – Section 4 – integrates and synthesises the evaluation results to provide an overview of the main evaluation conclusions and implications. Annex I provides a detailed breakdown of the scenarios' validation carried out in the Co-Labs. Annex II provides a detailed mapping of the digital challenges covered.

1.4 RELATION TO OTHER WPS & TASKS

As noted below in Section 2, evaluation is a cross-cutting activity in LIVE IT. It links to all project work-packages and activities.

2 EVALUATION APPROACH AND METHODOLOGY

2.1 THE 'OBJECT' OF EVALUATION

The 'object' of evaluation in this Report – the 'thing' that is being evaluated – is the LIVE IT project. As Figure 1 below shows, the central focus of the project is on designing and

implementing four 'Co-Labs' – co-design and co-production Laboratories - each of which is configured to the key settings in which people with cognitive disabilities engage with digital technologies in everyday life. Each of the four Labs has a distinct thematic focus: emergencies, self-development, life skills and health and support. The co-Labs provide spaces for teams of people with cognitive disabilities working in collaboration with experts and designers to explore and build on available platforms and tools to re-frame and reconfigure them to their needs. The co-Labs are supported by an online ***'Makerspace'*** in which people with cognitive disabilities can experiment with promising digital tools in a real-time virtual laboratory that can identify what works and what doesn't. The LIVE IT ***Cognitive Community*** connects an EU-wide network of design Labs – which includes people with cognitive disabilities – to the Makerspace platform.

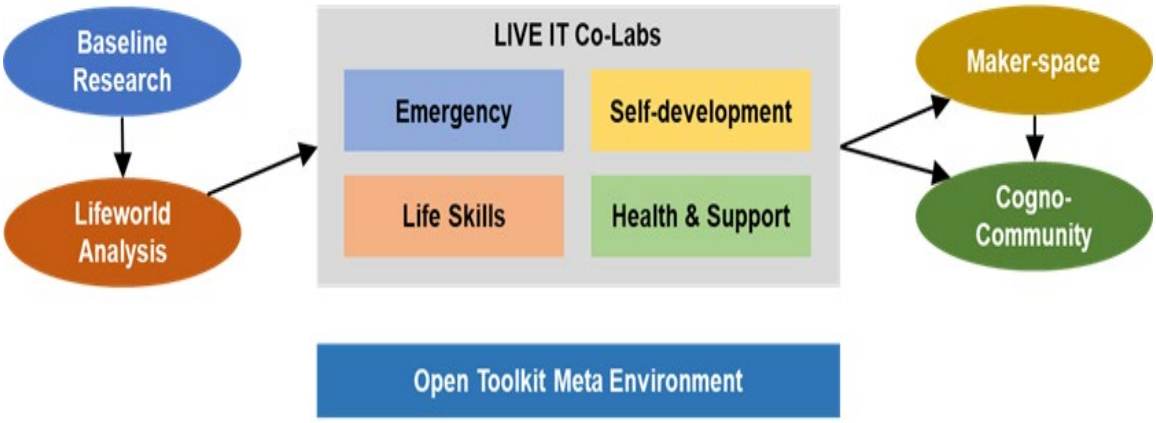

Figure 1 – LIVE IT approach

The results of the activities carried out in all of LIVE IT's work packages – in particular the collaborative experiments carried out in its 'Co-Labs' – feed into the production of the LIVE IT 'Open Toolkit' – which is the main output of the project. The Toolkit brings together all of the results and learning generated in the project and makes it useful for stakeholders by providing guidelines, tools and good practice examples to support web accessibility for people with cognitive disabilities. It consists of three sets of tools:

- 'Knowledge-gathering' tools. These cover: directories of cognitive disabilities, IT platforms, services and tools; a Directory of key people, organisations and places in the field
- 'Sense-making' tools. These cover: the LIVE IT 'Advisor Tool' - this offers users a selection of tools that have the potential to help people with cognitive disabilities meet their digital challenges, on the basis of their needs and profile
- Co-design tools. These cover: a 'Knowledge Community' to support knowledge sharing between the LIVE IT Community Labs and cognitive disability groups; an online 'Makerspace' that reinforces the co-design work being carried out in the Labs through supporting the broader community of people with cognitive disabilities to explore and experiment with the ideas and tools coming from the Labs, as well as A series of

'Hackathons' that focus on co-producing solutions to specific issues for people with cognitive disabilities ; Guidelines to support a range of stakeholders to deliver better web accessibility – from designers and developers of platforms and software tools through to policymakers with an interest in supporting digital inclusion agendas.

2.2 EVALUATION IMPLEMENTATION IN LIVE IT

Evaluation is a cross-cutting theme in LIVE IT. It is not covered by a single dedicated work package or activity but is delivered across four separate activities, as shown in Table 1.

Table 1 - Evaluation in LIVE IT

Activity	**Description**
3.2: Evaluation	Focuses on evaluation of the Open Toolkit, and forms part of the overall project evaluation, based on Theory of Change
3.3: Pilot testing	Focuses specifically on assessing the usability of the Open Toolkit
4.4: Outcome evaluation	Focuses specifically on evaluation of the Makerspace and Hackathons
5.3: Monitoring and Reporting	Design and implementation of internal project evaluation, including Quality Register, Process Dashboard and Partner Survey, feeding into overall project monitoring

In addition to these four activities, focusing on evaluation and pilot testing of the Open Toolkit, evaluation of the Makerspace and Hackathons and design and implementation of internal project evaluation, evaluation was also carried out in two other activities:

- Task 2.2: Scenario co-design – which entailed capture and analysis of data on how people with cognitive disabilities work with LIVE IT team members to experiment with scenarios in the Labs
- Task 2.4: Scenario Validation - which entailed capture and analysis of data to validate the scenarios developed in the Labs.

In order to cohere and integrate the results of these discrete activities, an over-arching evaluation framework was developed, which specified:

- the over-arching evaluation design
- the component evaluation activities that are incorporated into this over-arching design
- how these component activities relate to each other
- the evaluation purposes addressed by these component activities and the evaluation questions they address
- the methods and tools to be used to collect and analyse the data needed to answer these evaluation questions.

This is illustrated in Figure 2.

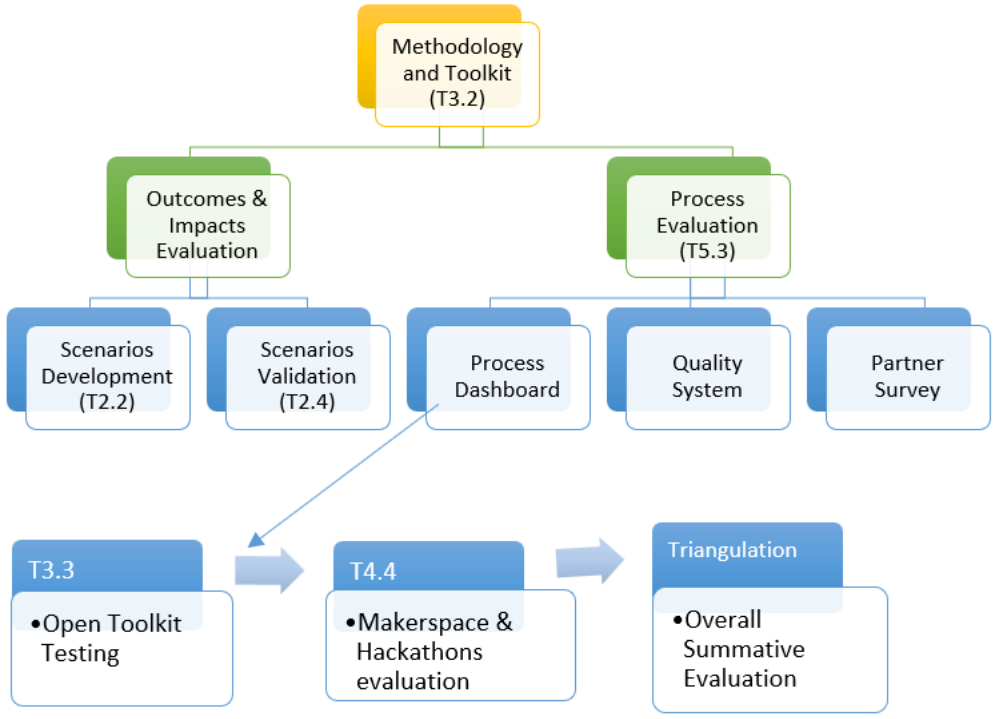

Figure 2 - LIVE IT Evaluation Methodology

As Figure 1 shows, the coherence and direction of the evaluation is shaped by the over-arching Evaluation Methodology and Toolkit. This is based on a 'theory-driven' evaluation approach, using 'Theory of Change', as described in the following section.

2.3 EVALUATION METHODOLOGY

The overall evaluation approach applied in the LIVE IT evaluation is based on an adaptation of the 'realist evaluation' approach (Pawson, 2006)[i]. This looks at how something is supposed to work, with the goal of finding out what strategies work for which people, in what circumstances, and how. A realist approach is theory-based and is essentially about testing a theory about what 'might cause change', even though that theory may not be explicit. One of the tasks of a realist evaluation is therefore to make the theories within an intervention explicit, by developing clear hypotheses about how, and for whom, programmes might 'work'. The implementation of the programme, and the evaluation of it, then tests those hypotheses.

A key tool in carrying out realist evaluation is 'Theory of Change'. Theory of Change tells the project 'story' – from the 'presenting problem' it addresses through to the change it hopes to make on that problem at the end of the project and beyond (i.e., the project's expected 'impacts'). Theory of Change thus gives us a framework for the evaluation because it makes the theories within the LIVE IT project explicit, by developing clear hypotheses about how, and for whom, LIVE IT might 'work'. The implementation of the project, and the evaluation of it, then tests those hypotheses and, if necessary, revises them in light of evaluation evidence. This means collecting data, not just about intervention impacts, but also the processes of the

intervention implementation. A simplified Theory of Change for LIVE IT is shown in the diagram below (Figure 3).

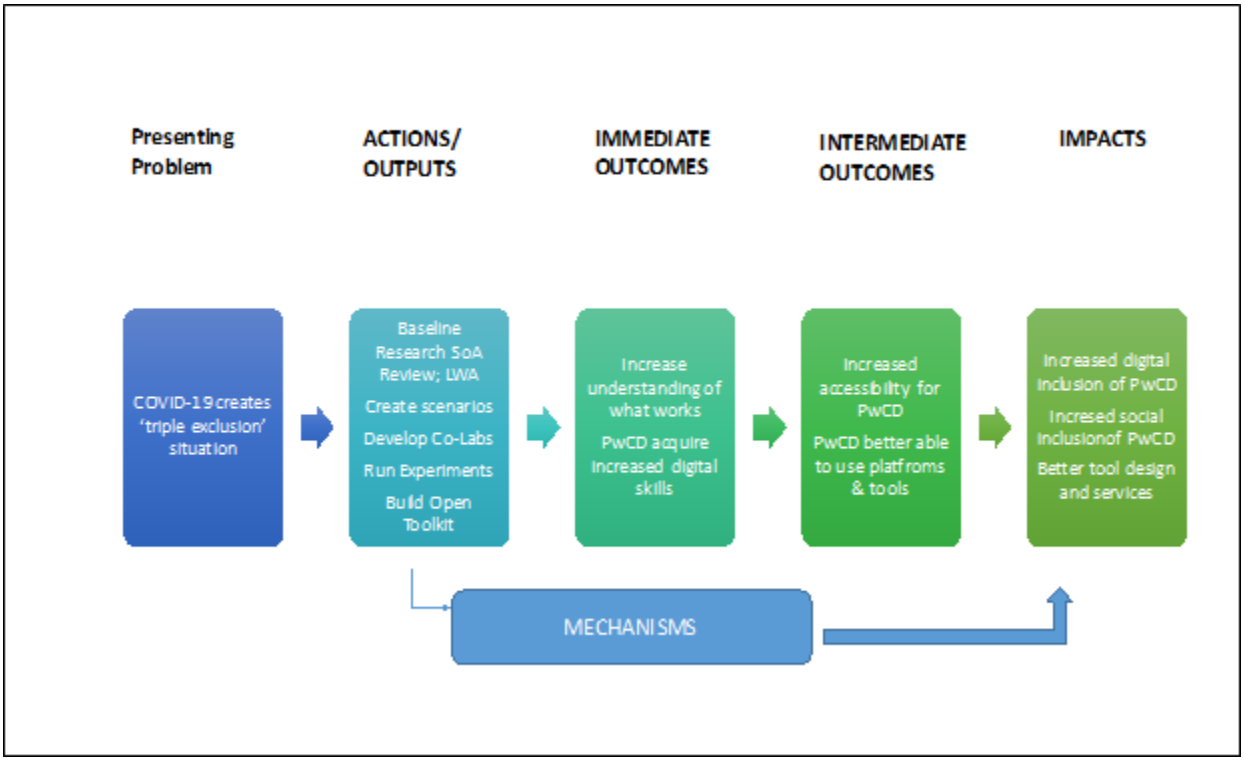

Figure 3 - LIVE IT – simplified Theory of Change

LIVE IT's Theory of Change maps the project 'change journey' so we can see the connections between the 'presenting problem' it wants to solve, the expected impact on that problem at project end and everything that's supposed to happen in between. It sets out the ***presenting problem*** the LIVE IT wants to address, i.e.:

- People with cognitive disabilities are excluded from digital life because of a combination of structural and digital inequalities, amplified by the migration online, post-COVID, of many aspects of everyday life, particularly services

and ends with the change the project wants to make to this problem after it has completed its journey – in other words the ***expected impacts*** LIVE IT will achieve at project end, i.e.:

- Better design of online content and services
- Increased digital inclusion of PwCD.

To get from presenting problem to expected impacts, LIVE IT carries out ***activities*** – for example carrying out a review of state of the art and lifeworld analysis.

These activities lead to the production of ***outputs*** such as the Co-labs set up in four different countries.

The utilization of these outputs leads to ***immediate outcomes*** (changes in awareness and increased knowledge), for example improved understandings of what works for the digital inclusion of people with cognitive disabilities.

These immediate outcomes lead to ***intermediate outcomes*** (changes in behaviour and structures), for example digital platform designers increasing the accessibility of the technologies they design.

Ultimately, these outcomes, combined together, will lead to the longer-term impacts LIVE IT aspires to. In particular theory of change looks to identify and understand 'mechanisms' – the combination of factors which operate in particular contexts to generate outcomes of interest (Befani, 2012)[ii].

LIVE IT's Theory of Change shapes all the evaluation activities implemented in the project and the methods and tools these activities apply.

The outcomes and impacts ('summative') evaluation can be seen as primarily linear in nature:

- It starts with capturing and assessing how our target group works with the project teams in the Co-Labs to explore different scenarios (Task 2.2)
- It progresses to validating the scenarios that are selected through the work of Task 2.2 as potentially providing maximum benefit to supporting the digital inclusion of people with cognitive disabilities (Task 2.4)
- The next stage of the outcomes and impacts evaluation is pilot testing the Open Toolkit (Task 3.3)
- This phase is followed by final evaluation of the Open Toolkit in the Hackathons delivered through the Co-Labs and in the broader on-line Makerspace (Task 4.4)
- The results of all the evaluation activities carried out are compared with each other and synthesised to produce a final summative evaluation of the project.

Each set of activities addresses different evaluation questions and uses different evaluation methods and tools. These are summarised in Table 2, which provides examples of the methods and tools used in the evaluation.

Table 2 - Evaluation Activities, Questions, Methods and Tools

Evaluation activity	**Key evaluation questions**	**Methods and Tools examples**
Process dashboard	Is the project meeting its milestones and targets?	Key Performance Indicators (KPIs) Partner Survey
Quality System	Are the project outputs of sufficient quality?	Peer and external review of deliverables
Scenarios development	Which kinds of technology platforms and tools increase	Observation analysis

Evaluation activity	Key evaluation questions	Methods and Tools examples
	the digital inclusion of people with cognitive disabilities?	Concurrent think aloud (CTA) analysis
Scenarios validation	Are the scenarios selected for inclusion in the Open Toolkit suitable for increasing the digital inclusion of people with cognitive disabilities?	Structured co-creation workshops Walk-throughs
Open Toolkit Testing	Does the Open toolkit conform to desired standards of usability, user-friendliness and user effectiveness?	Systems Usability Survey Walk-throughs Structured interviews
Makerspace and Hackathon Evaluation	How do users work with the Open Toolkit to co-design innovative solutions to key online challenges?	Observation Concurrent think aloud (CTA) analysis

The following section of this Report presents the results of the application of the methodology. We first look at how LIVE IT has performed set against key output targets. The report then covers the work carried out in the Co-Labs, with a particular focus on the participant experience and outcomes associated with participation. This is followed by an analysis of the Makerspace and Hackathon activities. We then present the results of the evaluation of the Open Toolkit. The section concludes with a look at the economic aspects of the evaluation.

3 EVALUATION RESULTS

3.1 OUTPUTS ANALYSIS

Discussion of the results of the evaluation begins with an analysis of the project outputs, set against key output targets as specified in the project DoA. These are summarised in Table 3, which shows the cumulative results of the 'process dashboard'. The process dashboard provides a set of indicators to assess progress against key project targets and milestones in relation to the project objectives.

Table 3 - LIVE IT Outputs Analysis

Objective	Indicators	Status at: 31/5/22	Project target
O1: Baseline Research	No. Literature review items and good practice cases intensively reviewed	180	50
O2: Lifeworld analysis	No. of LWA case studies and no. participants	4/150	4/60
O3: Review of platforms, tools and authoring systems	No. good practice cases of platforms, development tools and authoring systems assessed	233	100
O4: Set up pilot Labs in four locations – supported by a virtual Makerspace and online Community	No. LIVE IT Labs set up No. people participating in Makerspace and online Community	7 260	4 200
O5: Develop scenarios and test solutions to increase web accessibility for PwCD	No. users participating in Co-Labs	355	50
O6: Test Open Toolkit in Labs and Community through collaborative experiments	No. tools evaluated in Co-Labs and online Community	125	50
O7: Disseminate the learning from LIVE IT activities	No. visits to project website	1100	NS
	No. leaflets/newsletters distributed	5	NS
	No. in social media community	1412	1500
	No. webinars delivered	4	4
	No. local seminars delivered	4	4
	No. Conference presentations produced	3	2
	No. Peer-reviewed papers produced	3	2
	No. participants at final conference	250	100-200

NS = Target not specified in project DoA.

Table 3 shows:

- The 'research' activities carried out by LIVE IT (O1, O2, O3) have outperformed their targets. The target number of Literature review items and good practice cases

intensively reviewed has been exceeded by over 350%, and the target participants in the 'lifeworld analysis' has been exceeded by 250%.

- Over twice the target number of good practice cases of platforms, development tools and authoring systems were reviewed and assessed.
- LIVE IT set up and ran seven Co-Labs (O4), set against a target of four. The Co-Labs involved 355 people with cognitive disabilities in scenario validation experiments, exceeding the project target by a factor of seven (O5).
- 125 different tools were used in the Co-Lab experiments, exceeding the project target by 250% (O6).
- LIVE IT's outreach and dissemination activities present a mixed picture (O7). On the one hand, the project reached or exceeded its targets in terms of webinars and local seminars delivered, conference presentations and peer-reviewed journal articles produced and attendance at the final conference. However, LIVE IT had limited success in reaching its broader target audience of stakeholder communities.
- Web site visits – although not set a target in the DoA – have run at a low level. Social media community contacts recorded across all channels reached around 60% of the target. Although the target of stakeholders participating in the LIVE IT Community was exceeded by 17.5%, the level of involvement in the Makerspace was very low.

3.2 EVALUATION OF THE CO-LABS

The work involving people with cognitive disabilities as co-participants in exploring access challenges and as co-designers of solutions to these challenges took three main forms:

- identifying, mapping and assessing the online challenges people with cognitive disabilities face
- exploring and assessing the capabilities of a range of tools developed to increase web accessibility to meet these challenges
- co-developing solutions to the challenges.

3.2.1 Mapping and assessing digital challenges

The starting point for the work of the Co-Labs was creating a map of the challenges people with cognitive disabilities face online. This mapping drew on the preliminary research work carried out in work package 1 – in particular task 1.1, state of the art review and task 1.2, initiatives review – as well as the results of the lifeworld analysis carried out in Task 2.1, which aimed to understand the digital experiences and needs of people with cognitive disabilities from the perspective of their 'lived experience'. Building on this work, initial work in the Co-Labs combined two activities:

- mapping, categorising and clustering challenges, then exploring in-depth what these challenges mean for people with cognitive disabilities in practice, through focus groups, interviews and observational analysis of the use of various tool to address the challenges

- mapping challenges against the attributes of tools aimed at increasing web accessibility for people with cognitive disabilities in order to identify potentially promising tools and attributes to provide solutions to the challenges.

The challenges identified in the Co-Lab work were clustered according to the World Wide Web Consortium's (W3C) categorisation[iii] which documents user needs in terms of 'user stories' that highlight nine 'key topics' and story categories that depict actionable areas for developers/designers to improve web accessibility. Table 4 summarises the outcomes of the 'challenge mapping' exercise.

Table 4 - LIVE IT Challenges Mapping

Category	**No of tools mapped**
Memory (no of challenges 7)	136
Focus (no of challenges 7)	115
Time & organisation (no of challenges 12)	214
Mood (no of challenges 9)	157
Physical barriers (no of challenges 10)	137
Reading, writing comprehension (no of challenges 10)	158
Numbers (no of challenges 3)	15
Digital skills (no of challenges 19)	215
General (Developers) (no of challenges 7)	31
Total	**1178**

As Table 4 shows, just under 1,200 tools were identified, assessed and categorised. The analysis shows that the two most significant categories of challenge faced by people with cognitive disabilities, in terms of frequency of occurrence, are 'digital skills' –having the digital and media literacy competences to engage with the digital world – and 'time and organisation'- difficulties associated with dates, tasks and planning. In the 'digital skills' category the biggest challenges identified were navigating complex websites and problems associated with online learning. In the 'time and organisation' category the biggest challenges were difficulties sequencing and completing multi-step tasks and difficulties in planning for the future. Two other important challenge categories were 'reading, writing and comprehension', with the biggest problems covering difficulty in reading and writing, and difficulties following instructions – and 'mood', with the biggest problems covering difficulty handling the unexpected, anxiety, lack of motivation and handling emotions.

The next stage in this preliminary Co-Lab work was to map the challenges identified against the range of tools to support web accessibility for people with cognitive disabilities collated through LIVE IT's baseline research. The main objective of this exercise was to support the process of 'scenario co-design' and validation carried out in work package 2. This involved assessing how the tools collected and reviewed through the baseline research might offer solutions to the

needs and challenges PwCD experience in real life through the use of testable scenarios in the co-design labs. A 'scenario' is defined as a narrative description of the lifecycle of a user need within the LIVE IT project. A scenario tracks the user need from lifeworld analysis to co-lab activities to hackathon outcomes. A scenario includes descriptions of: problem(s) that need solution(s), platform(s) and tool(s) that could deliver a solution, and potential ways of using platform(s) and tool(s) in particular situations. To kick-start this work, LIVE IT carried out an analysis that mapped the attributes of the tools reviewed through its preliminary research in work package 1 to the challenges identified and analysed. Around tools were included in this mapping exercise. The results are illustrated in the graphic in Figure 4. Note this graphic is for illustrative purposes. It does not cover the full spectrum of tools used in the mapping exercise. This is provided in Annex II.

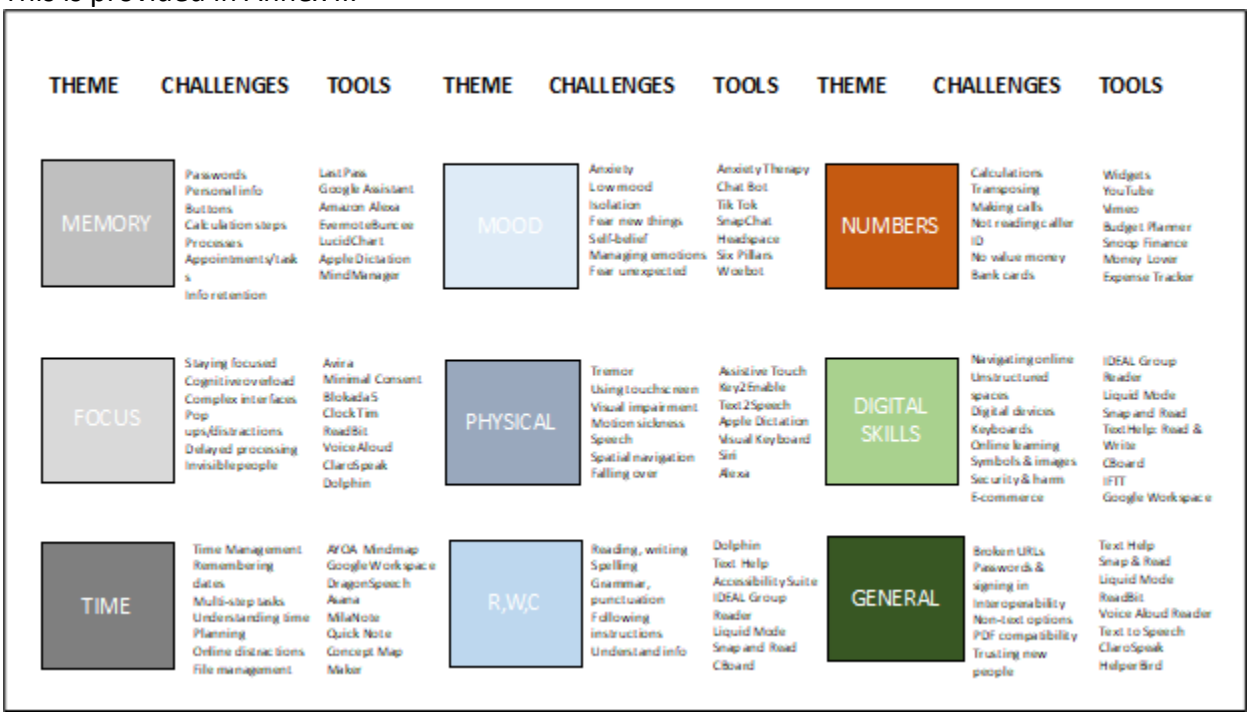

Figure 4 - Mapping tools against challenges

3.2.2 Co-Lab activities implemented

Using the mapping exercise as a foundation for subsequent exploration and experimental work in the Co-Labs, each Lab developed an implementation plan which specified the scenarios covered in the Lab and the experiments implemented within each scenario to explore and validate various combinations of tools that were expected to deliver improvements in web accessibility for people with cognitive disabilities. An experiment is defined as a testable hypothesis related to a particular scenario. Each experiment includes descriptions of procedures guided by a hypothesis, data collection tools, and a plan for data analysis.

Table 5 summarises the key headlines on the experimental activities carried out in the Labs.

Table 5 - Analysis of experimental activities in LIVE IT Co-Labs

Country	N. participants	N. events	N. scenarios	N. tasks	N. experiments	N. tools
Greece	23	23	8	19	132	21
UK	243	86	12	25	73	63
Ireland	25	18	7	14	50	21
Portugal	64	23	6	14	96	20
Total	**355**	**150**	**33**	**72**	**351**	**125**

Table 5 shows:

- A total of 33 scenarios were explored across the Co-Labs combined.
- A total of 72 tasks were carried out to explore these scenarios across the Co-Labs combined. 'Tasks' are bounded processes that define the actions participating people with cognitive disabilities are asked to carry out within the scenarios.
- A total of 150 experimental events were implemented to carry out these tasks.
- 355 participants in total were involved in these experimental events.
- A total of 125 different tools to support web accessibility for people with cognitive disabilities were validated in these experiments.
- The experimental activities were concentrated in the two UK Labs, in Bristol and London, and in the three Co-labs in Portugal. This reflects, firstly, the fact that multiple Co-Labs were set up in these two countries and, secondly, the catchment populations and profiles of these Labs, which were drawn from the wider community, and involved associate partner organisations providing a wide range of services in their communities.

3.2.3 Characteristics of Co-Lab participants

The profiles of participants involved in the Co-Lab work are summarised in Figures 5 to 7.

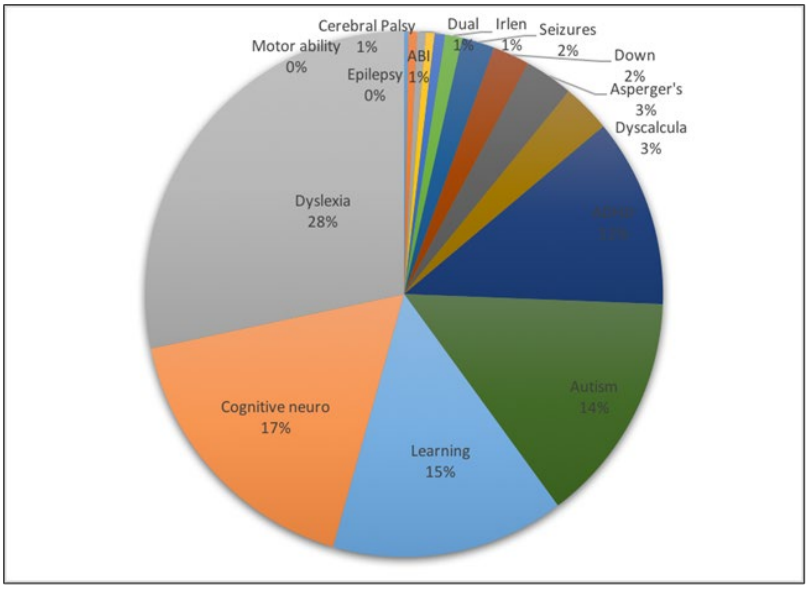

Figure 5 - Presenting cognitive disability issues reported by Co-Lab participants

Participants with cognitive disabilities in the Co-Labs show a broad spectrum of presenting cognitive disability issues. The majority – 28% - report dyslexia, 17% neurological issues and 15% each report learning difficulties and autism, with around 12% reporting ADHD issues. Other issues include dyscalculia, Asperger's and Down syndrome. A small proportion presented with severe cognitive challenges - cerebral palsy, epilepsy and motor neuron disease. Many participants presented with more than one cognitive issue.

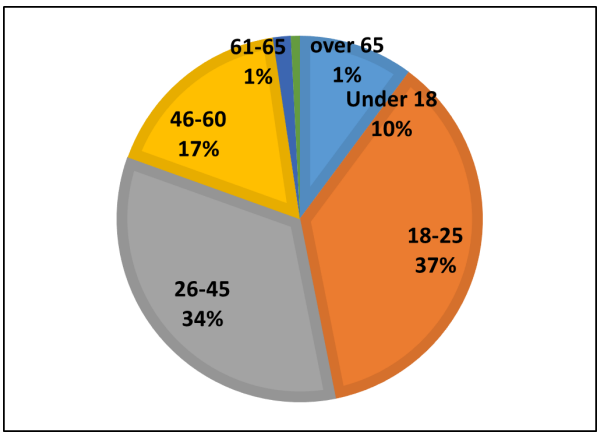

Figure 6 - Co-Lab participant age group

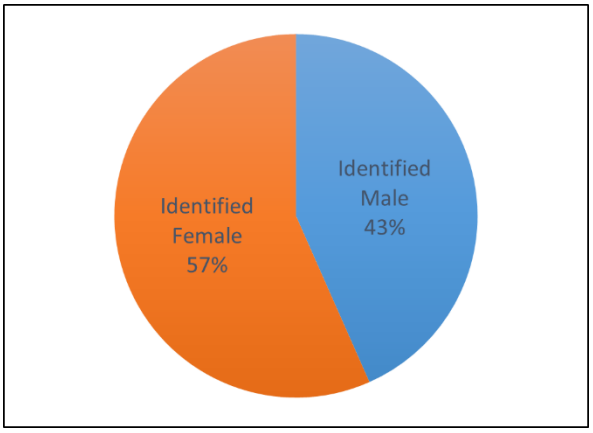

Figure 7 - Co-Lab gender distribution

As Figure 6 shows a broad spectrum of age was represented in the Co-Labs. Overall, young people 18-25 were proportionately the largest age group, with 37%. 34% of participants were in the 26-45 age group. 17% of participants were between 46 and 60 years; 2% were in the 61-65 age group and 1% above 65. 10% of participants were under 18. The age distribution varied considerably across countries. In Greece, 67% of participants were in the 18-25 age group and 45% of participants from Ireland in this age group. The Portuguese participants were even younger, with 53% in the 18-25 group and 47% in the under 18 age group.

As Figure 7 shows, overall, the gender distribution was slightly skewed towards females, with 57%, and 43% identifying as male. This distribution was reversed for Greece. In the UK and

Ireland, around 60% of participants identified as female. In Portugal, 69% of Co-Lab participants identified as male.

3.2.4 Co-creating solutions to digital challenges

As noted above each Lab developed an implementation plan which specified the scenarios covered in the Lab and the experiments implemented within each scenario to explore and validate various combinations of tools that were expected to deliver improvements in web accessibility for people with cognitive disabilities. A key objective of this work was to support people with cognitive disabilities to work with the LIVE IT researchers and experts in the Labs to co-design solutions to the challenges identified in LIVE IT's preliminary work. Essentially, the focus of the experimental work carried out in the Co-Labs was to apply co-design to meaningfully engage people with cognitive disabilities within 'validation scenarios' to produce web accessibility solutions. The task of scenario validation answers a series of questions: First, is a scenario doing what we want it to do? If yes, are there any ways the scenario can be improved for use within and beyond our co-design labs? If no, can the scenario be revised? Or, does a new scenario need to be developed?

The overarching methodology for scenarios development and validation mirrors that of the LIVE IT project as a whole in that it applies a 'design thinking approach' (Rowe 1987; Gobble, 2014; IDEO, 2014)[iv] [v] that involves: 1) an open exploration of the problem space (i.e. to develop the right things); 2) exploration of the solution space (i.e. to develop things right); and 3) the iterative alignment of these two phases whenever it is needed. This process is implemented through the co-creation process - empathizing – gaining an empathetic understanding of the problem; identifying – stating the problem from a human perspective; ideating – identifying new solutions by thinking outside the box; prototyping – developing new solutions to the problem; testing – evaluating the solutions.

Within this over-arching framework we used an adapted and abridged version of 'contribution analysis' to capture the processes and dynamics through which scenarios were explored and evolved. Contribution analysis is used in programme and project evaluation to create a causal chain – or 'contribution story' – that links actions and events to outcomes (Mayne, 2012)[vi]. In LIVE IT's scenarios development and validation, we wanted to capture the 'stories' for specific scenarios – how participants engaged with the scenarios; the actions they took to work with them and the mechanisms through which solutions to the challenges embedded in the scenarios were arrived at. In practical terms, this was done through a multi-methodological data collection approach, involving observation, walk-throughs, interviews and user feedback surveys.

Table 6 summarises the results of the scenarios validation in the Co-Labs. The Table shows, firstly, the scenarios covered in the initial implementation plan developed for each Lab; secondly, the number of scenarios that were revised as the scenario validation process unfolded and, thirdly, the number of new scenarios that were co-designed to address the issues identified through the validation process.

Table 6 - Scenarios Validation in the LIVE IT Co-Labs

Country/Scenario	Initial	Revised	New	Total
Greece	11	1	7	19
UK	16	7	2	25
Ireland	6	4	4	14
Portugal	10	2	2	14
Total	**43**	**14**	**15**	**72**

As Table 6 shows, 43 scenarios were validated initially in the Co-Labs. This involved a total of 280 different experiments. Of these, 191 were categorised as 'successful', defined as tools that successfully enabled participants to complete their assigned experimental task, and eliciting positive user feedback. In 92 of these experiments, 'pivots' were initiated. A pivot is where the experiment does not proceed as planned, because of problems experienced and identified in tool use, so that the experiment then proceeds along a different trajectory.

These pivots led to a total of 14 revised validation scenarios, involving 24 experiments. 5 of these experiments involved additional 'pivots' and delivered recommended changes to tool functionality. They also led to the development and implementation of 15 completely new scenarios involving 74 experiments.

Figure 8 shows the experimental 'success' rate for the Co-Labs. This is defined as the proportion of experiments that enabled participants to complete their assigned experimental task (successful) against the proportion of experiments that led to 'pivots', i.e., the development of revised and new scenarios and experiments.

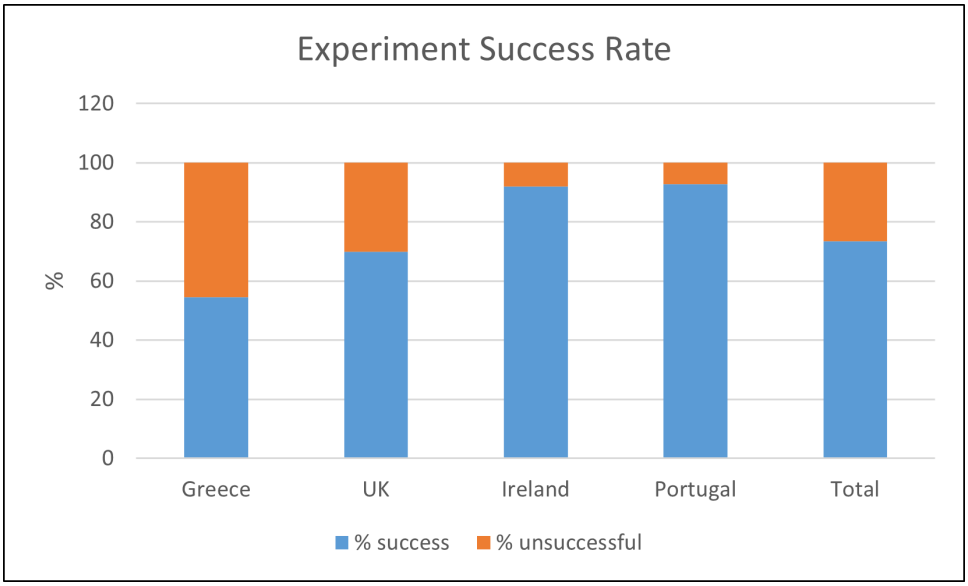

Figure 8 - Co-Labs Scenarios Validation Experiment Success Rates

The experiment success rate is a useful proxy for assessing the overall efficacy, usability and effectiveness of the range of tools aimed at increasing web accessibility for people with cognitive disabilities that was explored and evaluated in the Co-Labs. As Figure 8 shows, 74% of the 120 tools used in scenarios validation in the Co-Labs were classified as 'successful', i.e.,

tools that successfully enabled participants to complete their assigned experimental task and eliciting positive user feedback. The efficacy, usability and effectiveness of the tools used is discussed in further detail in Section 3.2.5 below.

Full coverage of the validation scenarios implemented across the Co-Labs is provided in an accompanying LIVE IT Deliverable – Task 2.4 Report. This is also summarised below in Annex I.

The scenario validation covered all nine of the topics specified in the WC3 framework (outlined above in Section 3.2.1). In addition, we mapped the challenges covered in the Co-Labs scenarios validation against the 'common accessibility challenges' identified in the "Inclusive Web Accessibility for persons with cognitive disabilities (Web Inclusiveness for All)" Call document that shapes the DoA on which LIVE IT is based, as shown in Table 7, which provides an estimation of the extent to which each challenge was addressed, shown as a relative proportion of the activities carried out overall.

Table 7 - Challenges covered in Co-Labs scenarios validation by Call document challenges

A. Challenge	**% coverage**
1. Signing in or completing authentication processes	5
2. Entering data or filling out forms, avoiding and correcting errors	10
3. Accomplishing multi-stage tasks	10
4. Using complex interfaces and intricate navigation patterns	8
5. Processing information on cluttered pages, avoiding cognitive overload	10
6. Getting focussed on the task at hand, limiting interruptions, reminders, and changes in content;	5
7. Stopping animations and auto-play videos or audios	3
8. Understanding text and graphics	10
9. Identifying key points from lengthy or heavy text-based pages	10
10. Interpreting and processing numerical information, including managing dates	3
11. Finding information	10
12. Following instructions	10
13. Looking for help or support	8
14. Using web authoring tools to create content	5

As Table 7 shows, all of the 14 accessibility challenges specified in the Call were covered in the Co-Labs scenarios validation. The challenges more comprehensively covered were challenge 2 - entering data or filling out forms, avoiding and correcting errors, challenge 3 - accomplishing multi-stage tasks, challenge 5 - processing information on cluttered pages, avoiding cognitive overload, challenge 8 - understanding text and graphics, challenge 9 - identifying key points from lengthy or heavy text-based pages, challenge 11 - finding information, and challenge 12 -

following instructions. Challenge 10 - interpreting and processing numerical information, including managing dates – was not well covered by the scenarios' validation. This reflects the relative lack of availability of tools in this category.

3.2.5 Tools evaluation

As noted above in the preceding sub-section 74% of the 125 tools used in scenarios validation in the Co-Labs were classified as 'successful', i.e., tools that successfully enabled participants to complete their assigned experimental task and eliciting positive user feedback. This sub-section provides more detail on the evaluation of the 125 tools for increasing web accessibility for people with cognitive disabilities that were used in the LIVE IT Co-Labs scenarios validation. To support monitoring and evaluation of the experiments carried out in the Labs an evaluation methodology and toolkit was developed that combined the following:

- An observation template
- A Survey template to collect data on user experiences with tools
- A 'journey mapping' tool to collect data on structured or unstructured walk-throughs with existing, apps/tools or technologies used in co-lab experiments
- Screen capture checklist
- Video, audio and image capture checklist.
- The ***Observation template*** is a structured checklist for recording the implementation of an experiment in the Co-Lab. It captures the experiment using six dimensions: the Environment in which the experiment takes place; the People involved; the Scenarios of use covered; the Objects deployed (tools); the Process followed, and the Outcomes recorded. The observation template can be used in two types of scenario – capturing the use of already existing tools or capturing the co-creation of a new tool from scratch.
- The ***Survey template*** is based on an abridged version of the item System Usability Scale (Brooke, 1986)[vii]. There are three scales to measure tool usability and usefulness (extent to which the tool/app meets user needs; ease of use of the tool/app; capacity of the tool/app to deliver the tasks required) and four 'semantic differential' scales to measure the user's 'emotional response' (overwhelmed-in control; distracted-focused; frustrated-encouraged; stressed-relaxed). Supplementary open-ended questions capture user perceptions of the challenges encountered using the tools; the things that worked well and the improvements needed.
- The ***journey map*** describes the steps the participant needs to take to accomplish the task/scenario covered in an experiment in the form of a journey or 'walk-through' (Klaar, 2014)[viii] Two kinds of walkthroughs are recorded: a structured walk-through takes the user through a pre-prepared sequence of steps and tasks. An unstructured walk-through records the user's 'natural journey' through an experiment.
- A final set of monitoring and evaluation tools enables real-time capture of user interactions with tools in Co-Lab experiments through automated screen capture and use of video and audio equipment to record experiments as they develop.

The journey mapping tool, screen capture checklist and video/audio recordings were used unevenly across the different Co-Labs, because of the different contexts, profiles, implementation systems and processes associated with each Lab. To support cross-Lab comparison, the data collected with these tools has not been included in the evaluation analysis presented below, although these data were used as input to develop the LIVE IT 'Open Toolkit'. The evaluation analysis is based on, firstly, overall scores for the main tools used in the experiments, measuring usability and usefulness together with user 'emotional response', as captured through the User Survey and, secondly, a summary of the key messages obtained through analysis of the supplementary open-ended questions in the User Survey that covered challenges encountered using the tools; the things that worked well and the improvements needed, together with analysis of the qualitative data obtained through the Observational analysis. This provides more detail on specific tools used in the scenario validation experiments and their usability and usefulness. Evaluation analysis for each Lab, or set of Labs, is reported on in turn for each of the four countries in LIVE IT. An overall aggregated evaluation assessment of the tools is then provided for the Co-Labs combined.

UK results

Lab Locations: London, Bristol
Lab theme: Emergencies
No. participants: 243
No. tools used: 63
Scenarios validated: 25, as follow:

- Finding help in an emergency (7)
- Preparing for an emergency (5)
- What to do in an emergency (3)
- Optimise phone for use (1)
- Participant directed (7)
- Managing healthcare (2)

Figure 9 shows how the tools used in the experiments are rated by users in the UK Labs in terms of Usefulness and Usability on the three evaluation criteria of ability to meet user needs, ease of use and capacity to support the implementation of the assigned task. The Figure shows the average score for each tool out of a maximum possible score of 5.

Figure 9 shows:

- User ratings varied on Usefulness and Usability across all of the tools explored in the UK Co-Labs, with an overall combined average score for all tools of 3.4/5 for 'meeting user needs', 3.7/5 for 'ease of use' and 3.6/5 for 'capacity to support the implementation of the assigned task'. All the tools bar two scored above the halfway mark on Usefulness and Usability, and five tools scored consistently high.
- The highest performing tools were 'Speechify', 'Immersive Reader', 'Sticky Notes' and 'ReadBit'.

- The lowest performing tools were ‘TalkBack’, ‘CBoard’ and ‘Voice to text - Android - Scenario 4’.

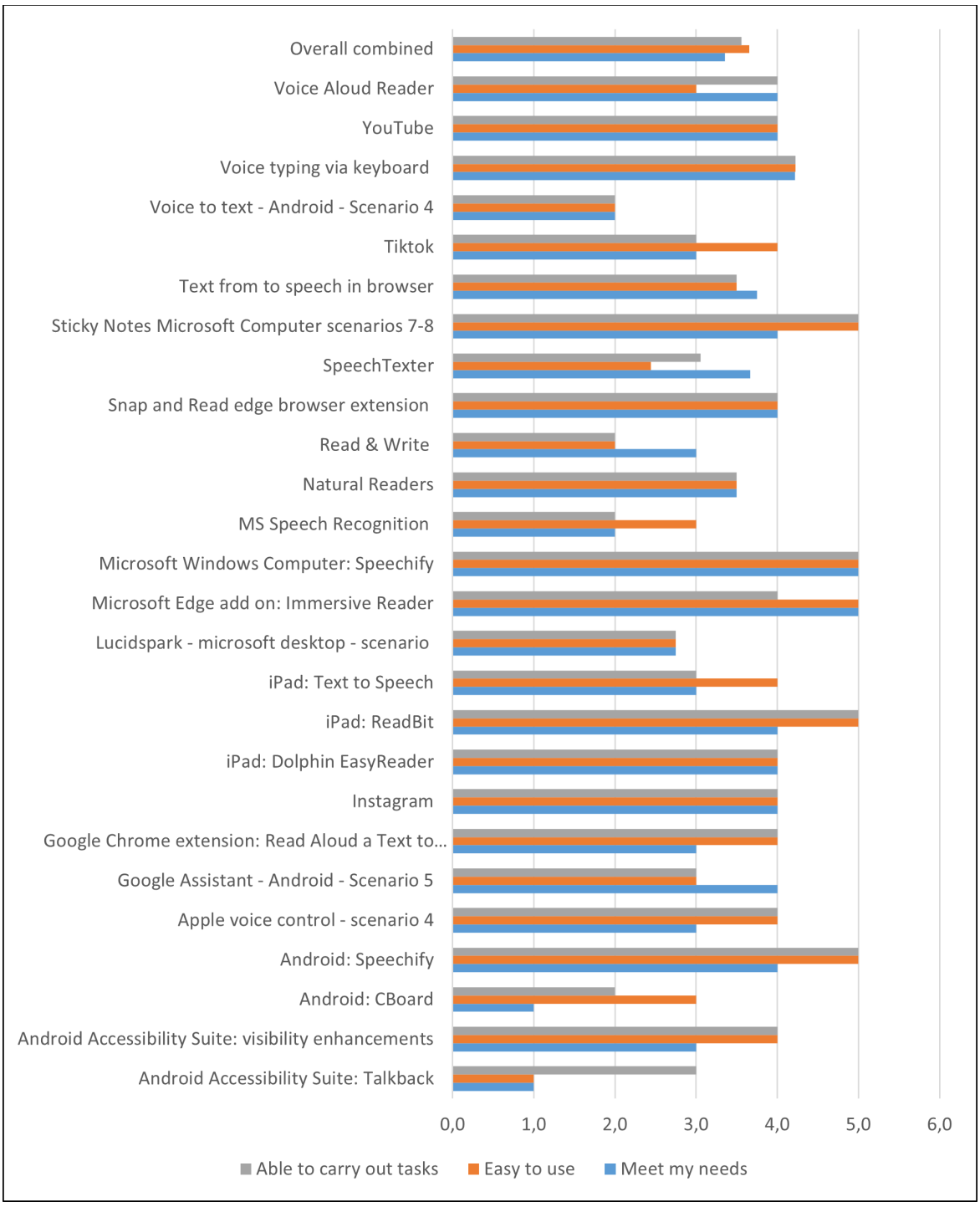

Figure 9 - Rating of tools on Usefulness and Usability, UK

Figure 10 shows how the tools used in the experiments are rated by users in in the UK Labs in terms of 'Emotional Response'. The four evaluation criteria use semantic differential (polar opposition) scales measuring overwhelmed-in control; distracted-focused; frustrated-encouraged; stressed-relaxed on a scale of -3 to +3.

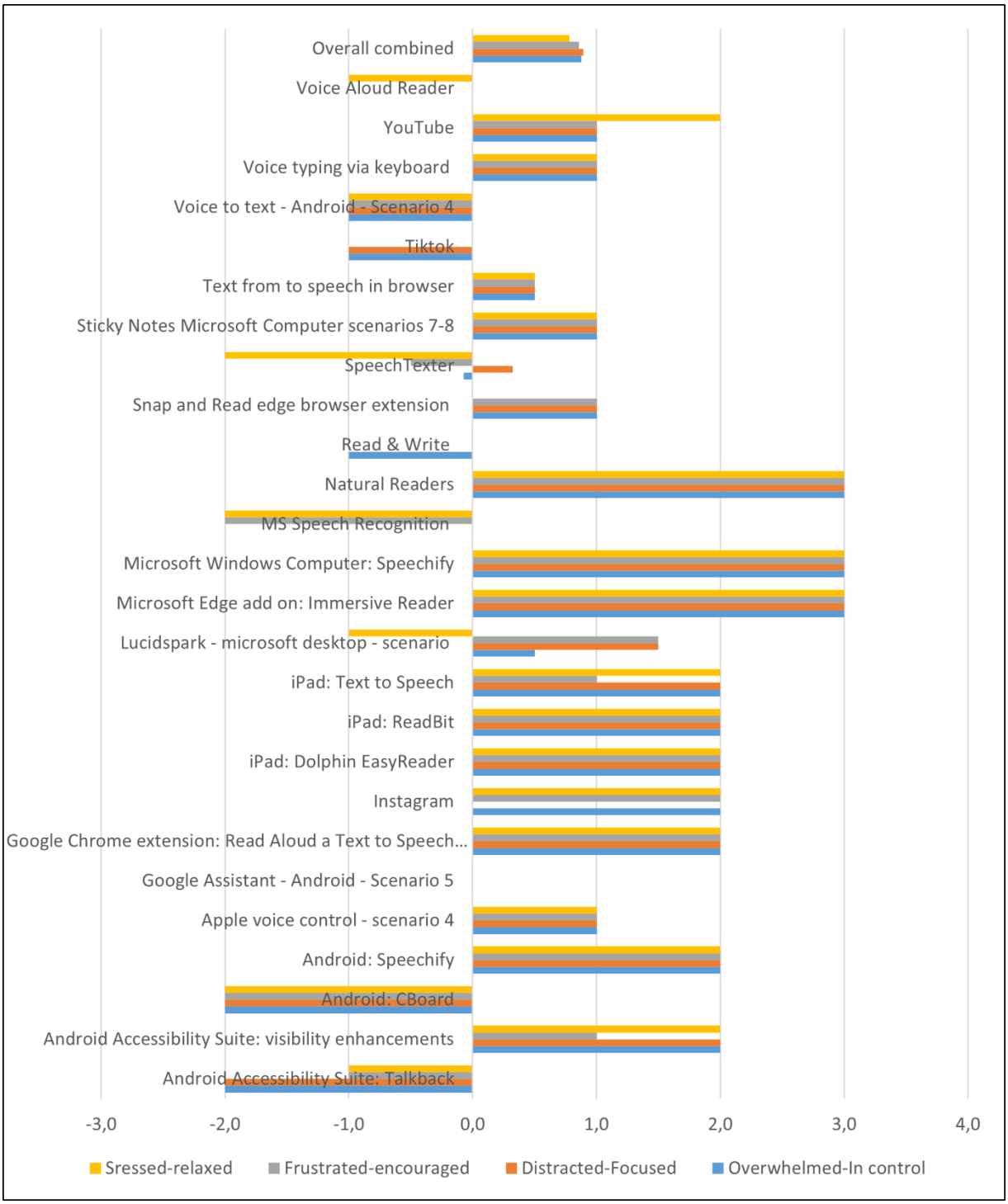

Figure 10 - Rating of tools on Emotional Response, UK

Figure 10 shows:

- Over 60% of the tools used in the UK Co-Lab experiments were rated above average on the 'Emotional Response' scales
- the highest performing tools on Emotional response were 'Immersive Reader', 'Speechify' and 'Natural Readers', each of which posted maximum scores.
- The lowest were 'Android: CBoard' with a total score of -4/12; 'Android Talkback', with a total score of -6, and 'MS Speech Recognition' and 'Voice to text - Android - Scenario 4', each with total scores of -4/12.
- The highest-performing tools overall were Immersive Reader' and 'Speechify' ', with a combined score on Usefulness, Usability and Emotional response of 26 and 27 respectively. Other high-scoring tools overall were ReadBit, with a total score of Dolphin EasyReader, with a score of 20.

Key messages from the Qualitative Analysis

- It has been difficult to implement formal scenarios work or stay strictly with emergency scenarios or 'emergency' as a category. Whilst this has been attempted in all cases, sessions have been allowed to progress organically - exploring internet and digital technology usage and challenges, as well as the way in which participants do engage with digital tech, tools and the internet in general. In many sessions this includes exploring what participants find useful and what an inclusive internet would look like.

- The reasons for this are numerous. Many of the participants are non-readers or have other significant behavioural and, or mental health challenges and learning difficulties. Many have significant financial issues and very limited access to digital technology. In Bristol the profile is slightly different but similar messages are coming across. Most of our participants so far suggest that audio-visual alternatives, simplified web summaries and signposting would be preferred. Some have suggested that chatbots are helpful, and others are clearly in need of extensive personalized toolkits in order to be able to effectively access the web as it stands without human support. There is a clear need for much more personalized support and training. The key message so far from the London and Bristol co-lab sessions has been how personal individual needs and challenges are.

- The LIVE IT toolkit, incorporating the results of the Co-Lab experiments, is seen as a significantly useful step forward. Some users do however seem to be suggesting that a fundamental re-structuring of the web, in order to be more inclusive, needs to include much more choice for users, providing perhaps audio-visual alternatives as an option. Text to speech apps, and voice search options are helpful, as are an array of other tools that address cognitive difference issues - such as overlays, mind maps, block selection and highlighting tools, font changes, visual identifiers and more.

- Our UK Co-lab work so far has shown that there are many more significant challenges to be overcome, particularly related to those that struggle with reading or can't read at all. Logging on, safety, security and trust issues also feature highly. Many of the co-lab participants are largely avoiding the text-based websites and relying largely on audio-visual sites such as YouTube and similar for information.

Greece Results

Lab Location: Aristotle University, Thessaloniki
Lab theme: Help and Support
No. participants: 23
No. tools used: 21
Scenarios validated: 19, as follows:

- Find and use contact information (5)
- Find a pharmacy (2)
- Navigate with Google maps (2)
- Obtain information via email exchange (2)
- Send message via email and social media (1)
- Find information online (3)
- Find COVID-19 information online and share with others (2)
- Organise and plan using online tools (2)

Figure 11 shows how the tools used in the experiments in the Aristotle University Lab are rated by users in terms of Usefulness and Usability.

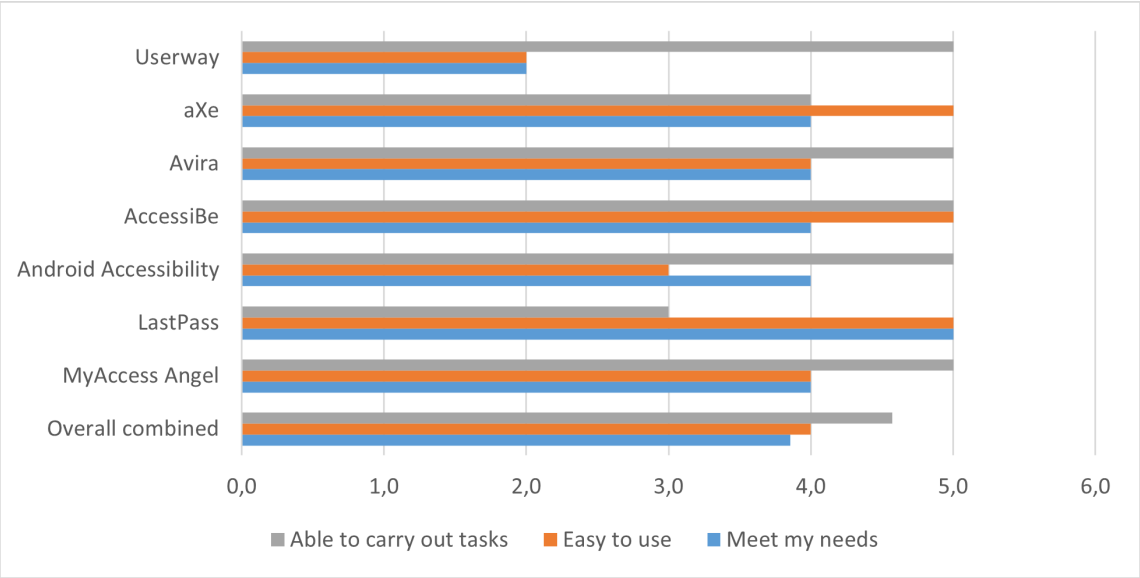

Figure 11 - Rating of tools on Usefulness and Usability, Greece

Figure 11 shows:

- User ratings were, overall, relatively high for Usefulness and Usability across all of the tools explored in the Thessaloniki Co-Lab, although the tools were rated lower on 'meeting user needs' – 3.9/5 – than 'ease of use' – 4.0/5 – and ''capacity to support the implementation of the assigned task' – 4.6/5.
- The lowest-performing tool was 'Userway', which averaged 2 (maximum score=3) on 'meeting user needs' and 'ease of use'.

- The highest-performing tool was ‘AccessiBe’, with combined score of 14 on Usefulness and Usability, with the other tools scoring 13, with the exception of ‘Android Access’ (12) and ‘Userway’ (9).

Figure 12 shows how the tools used in the experiments in the Thessaloniki Co-Lab are rated by users in terms of ‘Emotional Response’.

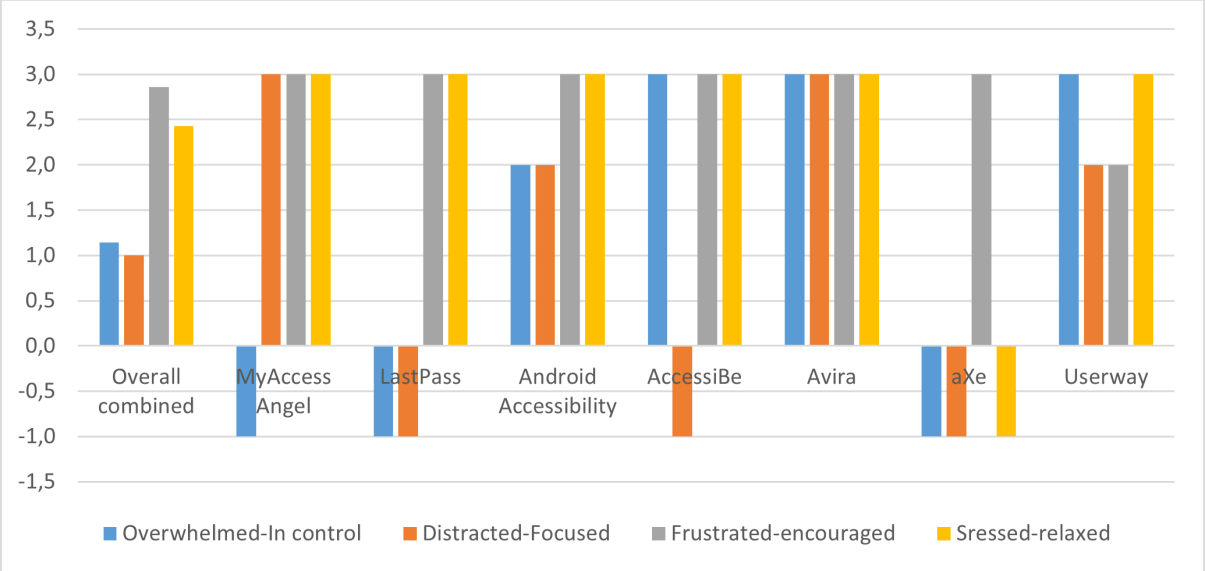

Figure 12 - Rating of tools on Emotional Response, Greece

Figure 12 shows:

- Overall, there was variability in scoring of the tools on ‘Emotional response’ by users in the Thessaloniki Co-Lab. Tools were scored relatively high on the ‘frustrated-encouraged’ and ‘stressed-relaxed’ poles, but lower on the other two criteria.
- The highest-performing tool on Emotional response was ‘Avira’, with a combined score of 12, followed by ‘Android Access’ (10).
- LastPass was the lowest performing tool, with a combined Emotional response score of 4.

The highest-performing tool overall was ‘Avira’, with a combined score on Usefulness, Usability and Emotional response of 25, followed by ‘Android Access’ and ‘AccessiBe’, with combined scores on Usefulness, Usability and Emotional response of 22 each.

Key messages from the Qualitative Analysis

- The tools used were considered potentially very useful, but some participants needed additional support and guidance from teachers and caregivers (particularly those with more severe cognitive disabilities).
- The main usability and usefulness issues highlighted by participants included language and translation issues, which impose barriers to comprehension, interaction, command reception; communicating without face-to-face contact, and the lack of guidance

provided for caregivers – a particular issue for participants who were unable to access the tools without caregiver support.

- Suggestions for improvements to the tools and software used in the *Help and Support* labs include: i) greater use of images and animations to better convey digital information to PwCD; ii) integration of voice user interaction capabilities, allowing users to employ speech commands to interact with the software; iii) options to improve opportunities for caregivers to support PwCD in the use of digital tools (e.g. guidance for caregivers to customise tools and apps specifically for individual needs).
-

Ireland Results

Lab Location: NUIG, Galway
Lab theme: Self-development
No. participants: 21
No. tools used: 25
Scenarios validated: 14

- Reading (2)
- Writing with dictation (2)
- Writing - typing (2)
- Reading & writing S2T (2)
- Reading & Writing (read-aloud feature) (2)
- Reading (finding text alternatives) (2)
- Reading (changing font) (2)

Figure 13 shows how the tools used in the experiments are rated by users in the NUIG Lab in terms of Usefulness and Usability on the three evaluation criteria of ability to meet user needs, ease of use and capacity to support the implementation of the assigned task.

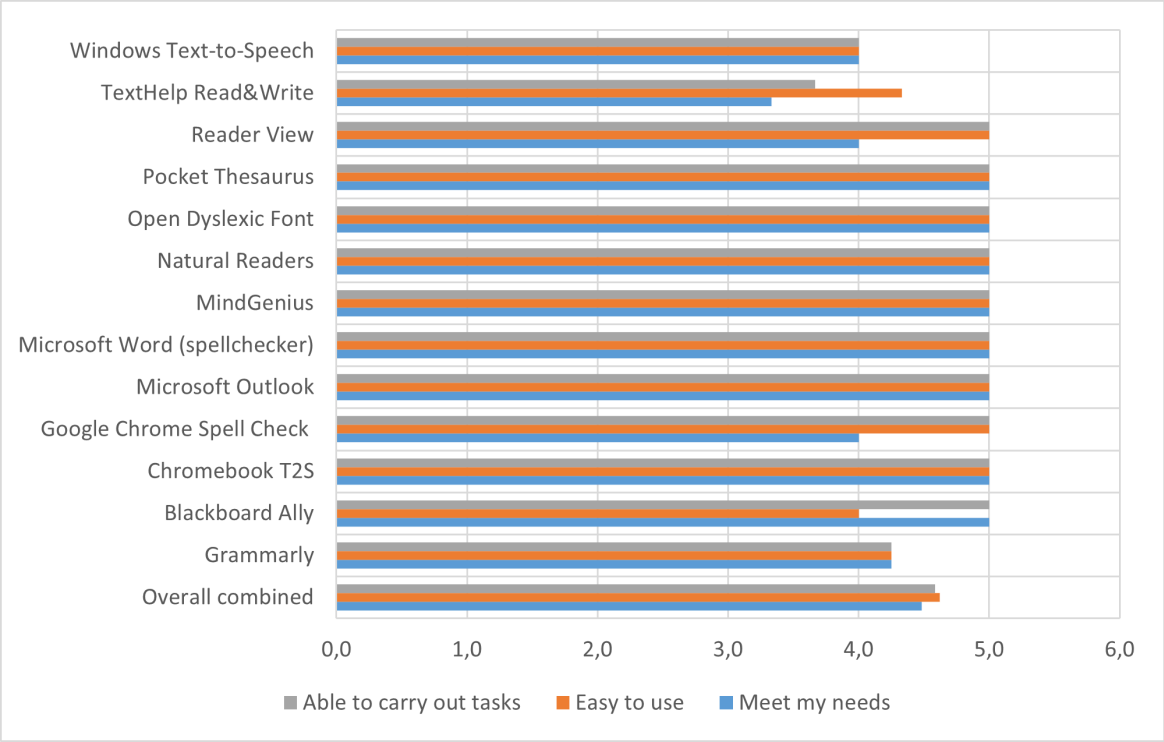

Figure 13 - Rating of tools on Usefulness and Usability, Ireland

Figure 13 shows:

- User ratings were consistently high for Usefulness and Usability across all of the tools explored in the NUIG Co-Lab, with an overall combined average score for all tools of 4.5/5 for 'meeting user needs', 4.6/5 for 'ease of use' and 4.6/5 for 'capacity to support the implementation of the assigned task'.
- All of the tools scored above 4/5 on the three evaluation criteria, with the exception of 'TextHelp Read&Write' which averaged 3.3/5 for 'meeting user needs', and 3.7/5 for 'capacity to support the implementation of the assigned task'.
- The highest performing tools were Chromebook T2S, Microsoft Outlook, Microsoft Word (spellchecker), MindGenius, Natural Readers, Open Dyslexic Font and Pocket Thesaurus.

Figure 14 shows how the tools used in the experiments are rated by users in the NUIG Lab terms of 'Emotional Response'.

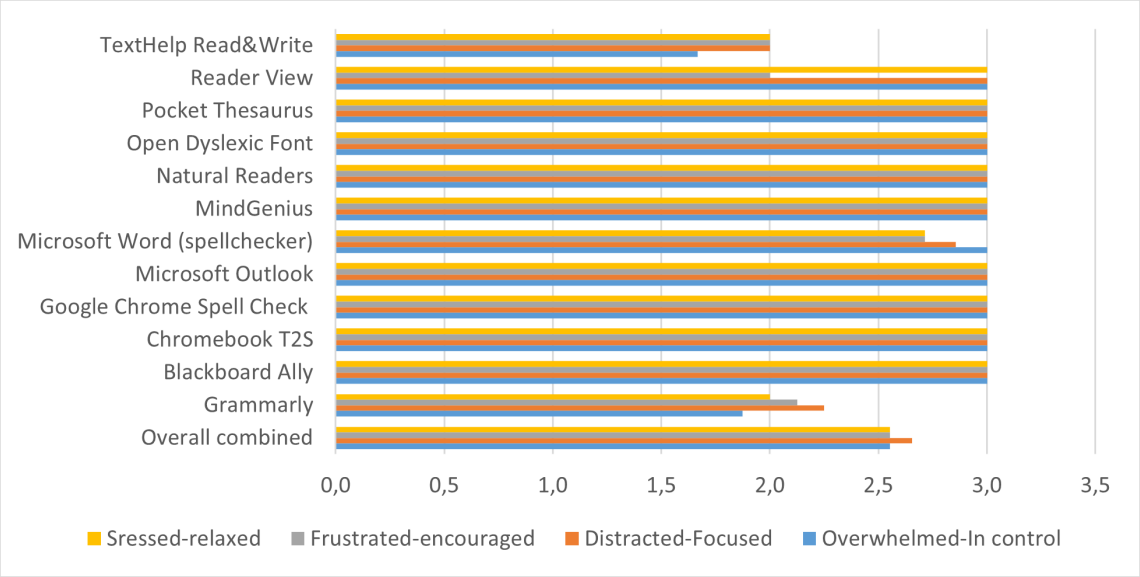

Figure 14 - Rating of tools on Emotional Response, Ireland

Figure 14 shows:

- User ratings were consistently high for Emotional Response across all of the tools explored in the NUIG Co-Lab, with an overall combined average score for all tools of 2.6 -2.7 for all four evaluation criteria (out of a maximum possible score of 3).
- As with the 'Usefulness and Usability' evaluation, the lowest performing tool was 'TextHelp Read&Write' which averaged 1.7 on the 'Overwhelmed-In control' scale and 2 for the other three scales. 'Grammarly' was also rated slightly lower than the other tools on Emotional response.
- As with the 'Usefulness and Usability' evaluation, the highest performing tools on the 'Emotional response' evaluation were Chromebook T2S, Microsoft Outlook, MindGenius, Natural Readers, Open Dyslexic Font and Pocket Thesaurus. Chromebook was also rated very highly on 'Emotional response'.

Several tools scored high ratings on overall performance – including Chromebook T2S, Microsoft Outlook, MindGenius, Natural Readers, Open Dyslexic Font and Pocket Thesaurus.

Key messages from the Qualitative Analysis

- The key usability issues highlighted for the 'Reading' scenario covered issues with ***text layout***, including problems with reading across multiple-column formats and non-horizontal text layouts; Issues with ***pausing, phrasing and stress;*** difficulties with ***processing mathematical equations;*** issues with software navigation; problems with aesthetic choices, for example, limited choice for voices to read texts and maintaining attention and focus on reading tasks.
- Suggested improvements proposed by users to Text-to-Speech tools and software for university students with cognitive disabilities include: improving text layout capabilities, providing more support for specialist texts (e.g., mathematical and scientific equations),

and providing more options for customisation, e.g., a wider range of voice and menu layout options.

- The key usability issues highlighted for the 'Writing' scenarios covered issues with dictation software not recognising different accents or user pronunciation errors; understanding complex terminology and understanding choices made by the software.
- Suggestions for improvements included supporting a wider range of accents and possible pronunciations and errors, providing more information around choices and better explanations to enable users to learn from mistakes. Additionally, tools could be 'smarter' in learning user behaviours and writing style, for example to stop suggesting changes that a user never accepts to avoid them having to reject the same changes every time. As one participant suggested the software should "just learn from me".

Portugal results

Lab Locations: UCP, Lisbon; Colégio Bola de Neve – Lisbon; ADFP, Miranda do Corvo
Lab theme: Life Skills
No. participants: 64
No. tools used: 20
Scenarios validated: 14, as follows:

- Text-to-speech (4)
- Translation (3)
- Complex tasks A (3)
- Complex tasks B (2)
- Understanding and interpreting numerical, text, and graphics information (1)
- Manage dates and appointments (1)

Figure 15 shows how the tools used in the experiments in the Portugal Labs are rated by users in terms of Usefulness and Usability.

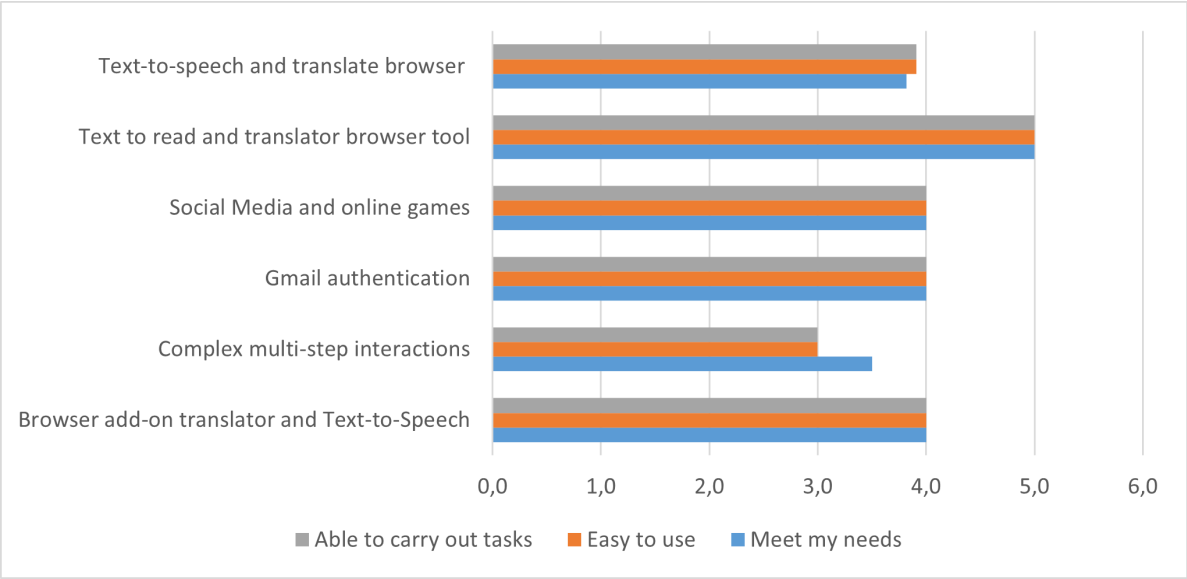

Figure 15 - Rating of tools on Usefulness and Usability, Portugal

Figure 15 shows:

- All tools with the exception of the tools to support Complex multi-step interactions were rated high on the usefulness and usability scales.
- The highest performing tools were the Text to read and translator browser tool followed by the Browser add-on translator and Text-to-Speech tool, Gmail authentication tool and Social Media and online games.

Figure 16 shows how the tools used in the experiments in the Portugal Co-Labs are rated by users in terms of 'Emotional Response'.

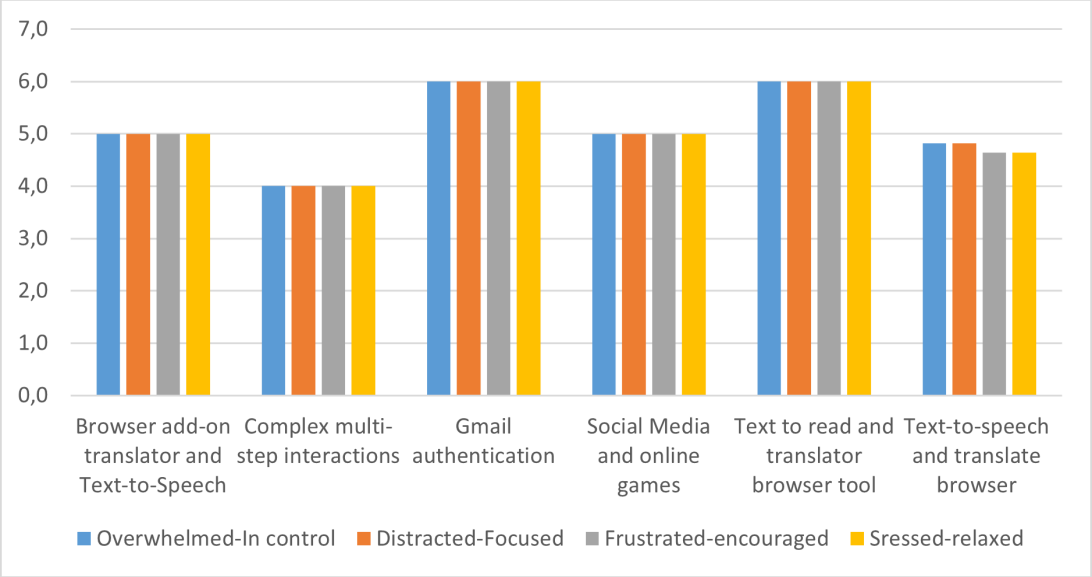

Figure 16 - Rating of tools on Emotional Response, Portugal

Figure 16 shows:

- User ratings for emotional response are similar to the user ratings for usability and usefulness, with highest performing tools the Text to read and translator browser tool followed by the Browser add-on translator and Text-to-Speech tool, Gmail authentication tool and Social Media and online games.

Overall, the highest-performing tool was the Text to read and translator browser tool, with a combined score on Usefulness, Usability and Emotional response of 39, followed by the Gmail authentication tool, a with combined score on Usefulness, Usability and Emotional response of 36.

Key messages from the Qualitative Analysis

- The text to speech tools used were considered potentially very useful in supporting users with reading challenges – in one case involving a user who could not read at all.
- However, the text to speech tools presented some challenges in terms of usability. For example, most users had problems selecting specific parts of a text that required

conversion to speech.

- Users proposed simplifications to the tools, with the end goal of enabling text conversion through a 'one click' button.
- The language translator tool was considered to be extremely useful, particularly in association with text to speech tools.
- Complex tasks, as their name suggests, proved complex for users to implement. The experiment with obtaining the COVID 19 vaccination certificate underlined a need for human support – through intermediaries – to assist users through the process of task delivery and completion and to explain what is needed at each step of the process.
- The tasks involving authentication and signing in were shown to be particularly challenging. In the case of the 'password manager' tool users found themselves in a 'Catch-22' situation in which they needed to create and memorise a password in order to use the tool. Users therefore prefer 'traditional' methods to address authentication and signing in challenges, such as asking for help from family members, teachers, friends, writing down passwords on paper and creating new accounts when necessary.

Overall Results

Although different tools were used in different contexts across the Co-Labs within different validation scenarios, combining results across all of the tools and the Labs to produce an aggregate 'meta-evaluation' score for the tools as a whole is a useful indicator of the extent to which the spectrum of tools available to increase web accessibility for people with cognitive disabilities.

Figure 17 shows the aggregate meta-scores for the tools on the 'usability and usefulness' evaluation measures.

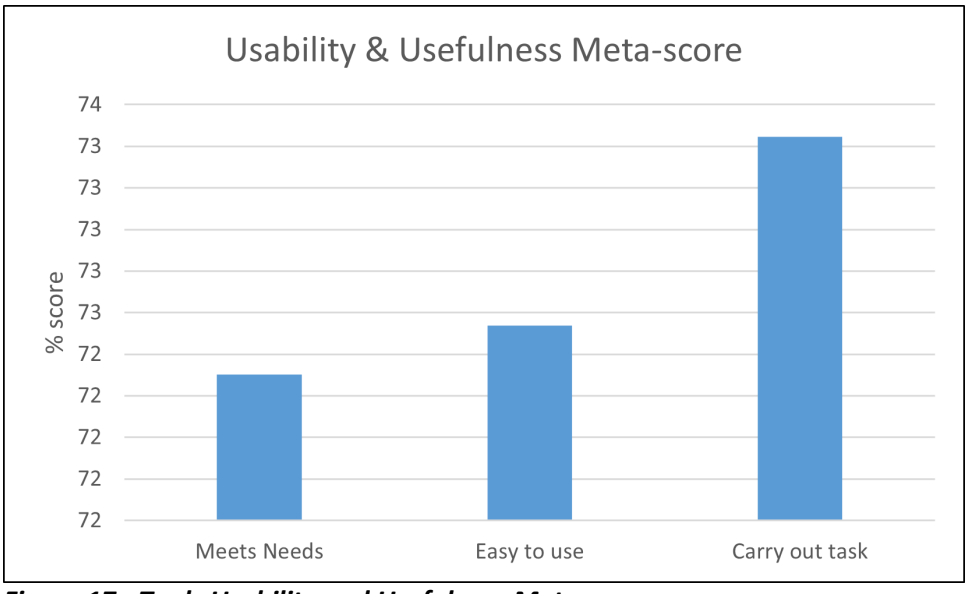

Figure 17 - Tools Usability and Usefulness Meta-scores

As Figure 17 shows the aggregate meta-score for the tools on the 'meeting user needs' measure was 72.3, out of a possible maximum of 100, with scores of 72.5 for the 'ease of use' measure and 73.4 for the 'task-enabling' measure. These scores suggest a relatively high level of usability and usefulness for the tools evaluated in the Co-Lab experiments.

Figure 18 shows the aggregate meta-scores for the tools on the 'emotional response' evaluation measures.

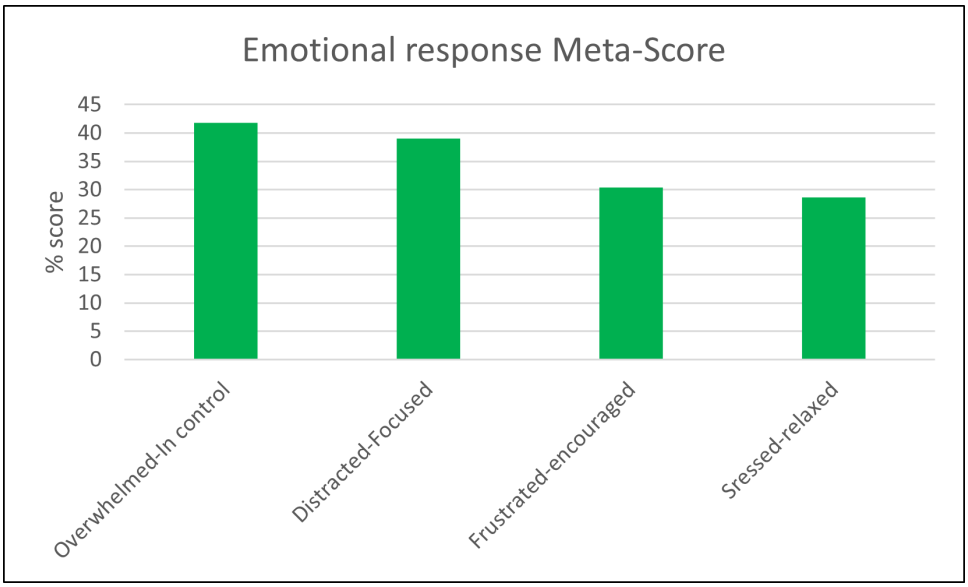

Figure 18 - Tools Emotional Response Meta-scores

As Figure 18 shows, although the tools evaluated in the Co-Labs scored relatively high on usability and usefulness, aggregate score on the 'emotional response' measures were much lower, with score of below 50% recorded on all four measures, with the lowest aggregate score – 29 – recorded on the 'stressed-relaxed' measure. This suggests that, whilst the range of tools currently available to increase web accessibility for people with cognitive disabilities perform relatively well on 'functional dimensions', there is a long way to go in terms of tool design in order to meet users' emotional needs. These results resonate with the data collected through LIVE IT on the challenges people with cognitive disabilities face in the digital world, which highlighted the sense of powerlessness felt by users of digital tools and services; the extent to which they perceived that many web-sites were designed to distract users; the sense of frustration felt at needs not being met as a result of the poor functionality of digital platforms, tools and services and the incipient anxieties generated when engaging with the digital world.

3.3 EVALUATION OF THE MAKERSPACE AND HACKATHONS

As noted above in Section 2.1, the LIVE IT Makerspace and its Hackathon programme are key to reinforcing and expanding the work implemented in the Co-Labs. The work carried out in the Makerspace is covered in detail in Deliverable 4.2 and the Hackathons are reported on in Deliverable 4.3. Presented below are the key messages from the evaluation of the Makerspace and Hackathons.

Makerspace

The Makerspace is intended as an online environment in which the co-design experiments carried out in the LIVE IT 'Co-Labs', aimed at developing web accessibility solutions for people with cognitive disabilities, are supported more broadly and widely. It provides a more open and widely accessible environment in which co-creation workshops can be used to experiment and test potential platforms and tools. This reinforces the co-design work being carried out in the Labs through supporting the broader community of people with cognitive disabilities to explore and experiment with the ideas and tools coming from the Labs. Within the Makerspace, people with cognitive disabilities can experiment with promising digital tools in a real-time virtual laboratory that can identify what works and what doesn't.

The Makerspace was developed using the 'Circle.so' platform. It has three 'areas'. The LIVE IT Community welcome area includes two sections, a 'Start Here' section that introduces new members to the LIVE IT project and a Discussion Space for facilitating member posts and discussions. The LIVE IT DevSpace is in three sections: Tools, which introduces useful tools and resources for people with cognitive disabilities, carers and family members, as well as researchers; Scenarios - dedicated to LIVE IT project scenarios from the co-labs, with more information about the co-design and validation of scenarios during the project; Developers, which provides useful tools and resources for developers, offering useful findings and guidance from the project. The LIVE IT Challenges area includes resources that aim to address four of the W3C challenge topics.

Evaluation of the Makerspace combined three sets of activities and evaluation methods:

- Analysis of membership levels and usage data
- Analysis of Makerspace activities
- Collection and analysis of feedback data on the Makerspace experience, through a user survey.

Analysis of membership levels and usage data show that engagement in the Makerspace is significantly lower than expected. The Makerspace platform currently has just over 30 members, and its profile is largely made up of LIVE IT researchers and co-lab members. These data show, on the one hand, a low level of use and, on the other, little reach beyond the LIVE IT project consortium, set against the key project objective of deploying the Makerspace as a platform to expand LIVE IT's sphere of influence to a broad range of stakeholder constituencies.

In terms of activities, the Makerspace currently has a total of 58 posts across all 10 spaces. Most posts have at least one like from community members, but very few posts have comments. The infographics in the Challenges area are the most popular posts with 10 likes and comments. Most posts were instigated by LIVE IT members, with only a few forum posts started by those outside of the central project team. Again, these activity patterns are considerably below expectations.

User feedback on the Makerspace was generally positive, though based on a small sample, with average ratings for usefulness, and above average for usability and navigation. A majority of users reported they would continue to use the Makerspace. Suggestions for improvement included adding more contrast, starting more discussion groups, providing better links to

relevant resources and providing more disability-specific information (e.g., specific tools for ADHD), as well as information on how to connect with other participants (e.g., other parents of children with ASD).

The evaluation shows that, despite the technical achievements of the Makerspace, and its largely favourable user response, it has struggled to attract and retain members. This is probably due to a number of factors, including the lack of a co-ordinated social media strategy for recruitment; the prolonged development and testing period needed to get the platform operational and the 'episodic' nature of content development and adjustment, which saw content constantly being re-assessed based on findings from the co-labs and hackathons. The latter meant that spaces in the Makerspace were significantly altered, and this could potentially have been disorientating for participants. Future changes to support improved recruitment and retention could include better, more targeted on-boarding at the beginning of member journeys, better signposting for how to use the space and the introduction of additional community features, such as polls.

Hackathons

Like the Makerspace the Hackathon element of LIVE IT was designed to engage with key stakeholder communities in order to: raise awareness of LIVE IT and, more broadly, web accessibility and design for all issues; involve these key stakeholder communities in sharing experiences and ideas; involve them specifically in contributing to the co-design of accessibility solutions. They were particularly aimed at, firstly, supporting the experimental work carried out in the Co-Labs by exploring web accessibility solutions for people with cognitive disabilities and, secondly, providing a space to support the development and testing of the LIVE IT Open Toolkit.

Five Hackathons were implemented in the Hackathon programme. Each was focused on a specific theme, set to reflect the key digital challenges for people with cognitive disabilities as identified in the literature and highlighted in the PPAG Call document. At the same time, the Hackathons were a primary source for collecting and analysing user feedback on the Open Toolkit as it developed. From the first Hackathon onwards, all of the updates, upgrades and improvements made to the Open Toolkit, were based upon feedback collected through the Hackathons.

Evaluation of the Hackathons combined four sets of activities and evaluation methods:

- Analysis of participation rates and participant profiles
- Monitoring and documentation of the activities carried out and their outputs, through observation, Toolkit walkthroughs and structured interviews
- Collection and analysis of user feedback data on the tools used and on the Open Toolkit, through surveys and individual and group interviews
- Collection and analysis of feedback data on the participant experience, through individual and group interviews.

Analysis of participation data shows a high level of engagement, with a total of 200 participants in the five Hackathons combined. As Table 8 shows, these participants made up a total of 38 different teams.

Table 8 - Hackathon participation

Hackathon	N. teams
1 Facilitating authentication and online forms	6
2 Processing heavy text-based and cluttered pages	9
3 Understanding and interpreting numerical, text and graphics information	12
4 Handling complex interfaces, multi-stage tasks and limiting interruptions	4
5 Accessible support and instructions	7
Total	**38**

The participants and teams were largely representative of the project partner countries – Greece, Ireland, Portugal and the UK, but also included participants from the wider stakeholder community, primarily as a result of the involvement of 'Design for All Europe'.

The Hackathons main focus was on experimental activities, mirroring those carried out in the Co-Labs. These experiments covered two sets of tasks. Firstly, participants explored a range of tools aimed at supporting web accessibility, according to the particular theme covered in each Hackathon. Second, they worked with the evolving LIVE IT Toolkit to assess its usability, usefulness and effectiveness. The outputs of these activities took the form of 'interventions' – defined as recommendations for tool and Open Toolkit improvement, ideas for new tools to address a perceived gap in user needs, or specifications for new tools. These interventions were presented in various formats, including text, images, PowerPoint presentations, screenshots.

Table 9 provides an indication of the level of activity and outputs generated in the Hackathons. As Table 9 shows a total of just under 200 interventions were created by participants in the Hackathons. With regard to the outputs specifically intended to improve the Open Toolkit, most of the proposals made by participating teams have already been implemented or are in the process of being implemented in the platform, making the toolkit more accessible, more functional and more user-friendly, and closer to end-user needs.

Table 9 - Hackathon outputs

Hackathon	N. interventions
1 Facilitating authentication and online forms	35
2 Processing heavy text-based and cluttered pages	58
3 Understanding and interpreting numerical, text and graphics information	54
4 Handling complex interfaces, multi-stage tasks and limiting interruptions	31
5 Accessible support and instructions	21
Total	**199**

Analysis of these interventions showed they could be classified into three categories:

- 'positive' interventions that highlight the value of a particular tool or an aspect of the Open Toolkit
- 'neutral' interventions that validate a particular tool or an aspect of the Open Toolkit in terms of usefulness, usability and effectiveness, and which may be accompanied by suggestions for improvement
- 'negative' interventions that document unsuccessful interactions of users with a particular tool or an aspect of the Open Toolkit and which indicate a significant failure to address user needs.

Figure 19 shows the results of the analysis.

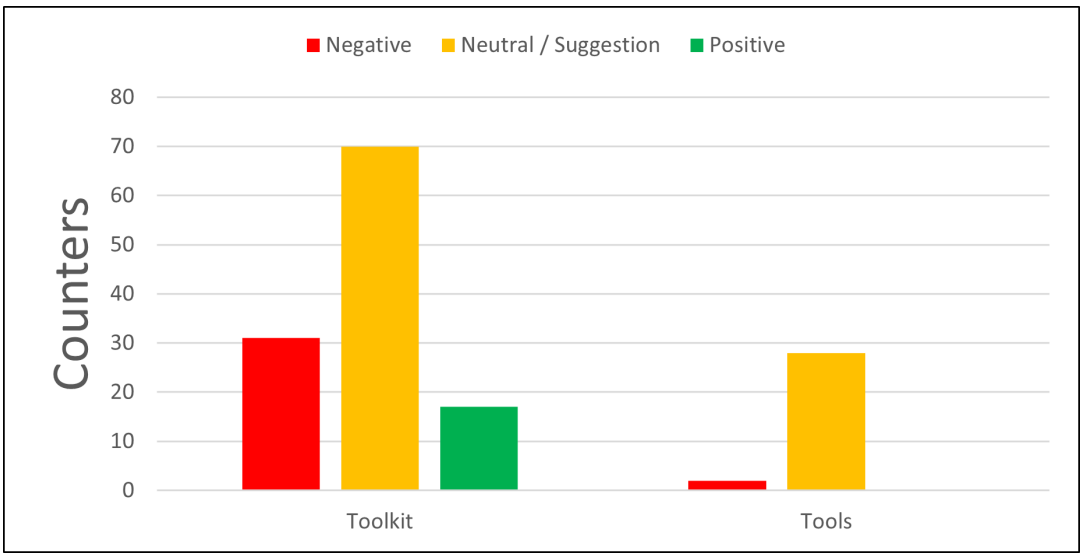

Figure 19 - Hackathon participant feedback on tools and Open Toolkit

As Figure 19 shows the majority of interventions in the Hackathons relate to the Open Toolkit. Most of these could be categorised as 'neutral' – validating the Toolkit whilst at the same time offering suggestions for improvement. However, a significant number of interventions – 30 – reflect negative user feedback. With regard to the interventions generated as a result of exploration with a range of tools being used in the Co-Labs, the majority can be categorised as 'neutral' with a small number categorised as 'negative'.

Further analysis of the evaluation data collected through the Hackathons allowed categorisation of the interventions generated on the basis of clusters of common content. Seven clusters were identified:

- Compatibility – the extent to which different platforms or operating systems support interoperability
- Content – aspects of tools or the Open Toolkit that would require some adjustment of content Exploitation - ideas for future use and ways to exploit the Open Toolkit or another tool
- Function – interventions that refer to the adjustment of an existing function or the addition of an additional function that solves a problem
- Language - issues created because of language, usually because a tool is offered only in an English version and not in local languages
- Layout – interventions related to fonts, graphics, alignments, sizes, colours and presentation of digital content
- Usability – interventions regarding the ease or difficulty in operating a tool and ways to enhance or facilitate this interactive digital experience.

Figure 20 shows the distribution of Hackathon interventions by each cluster.

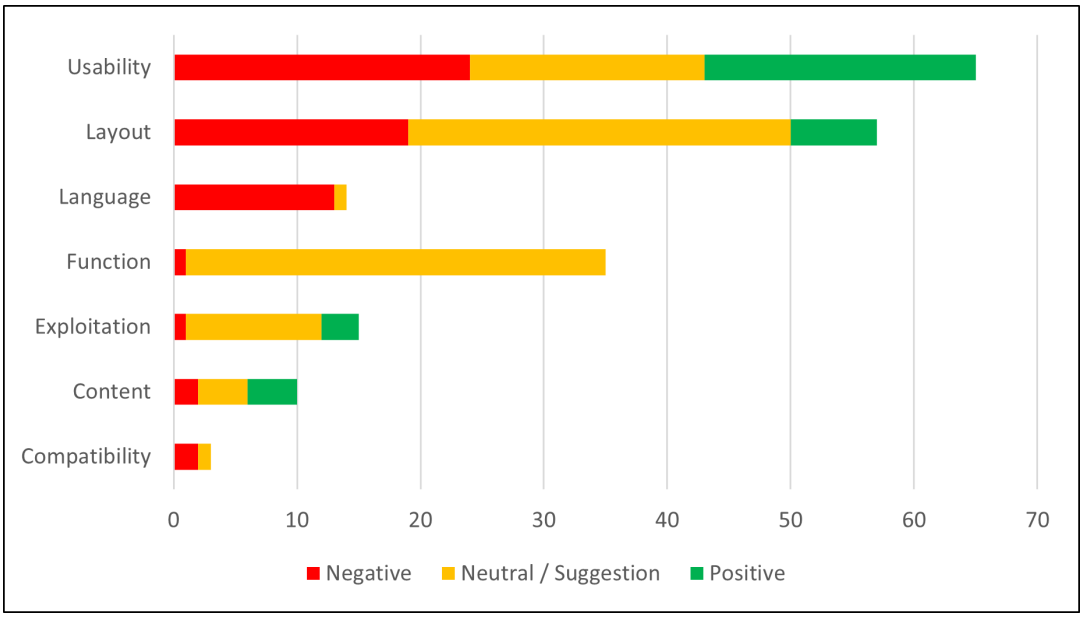

Figure 20 - Hackathon interventions by thematic cluster

As Figure 20 shows, the majority of Hackathon interventions focus on 'usability' issues. These were broadly split between positive, neutral and negative positions, although a significant number of negative usability issues were documented. The second biggest category of interventions focus on 'layout'. The majority of these interventions are classified as neutral, but a significant number of negative layout issues were documented. Language emerges as a major problem, as reflected in the large proportion of negative interventions in this category.

Language was also highlighted as a challenge in terms of evaluation of the Hackathon participation experience. Participant feedback on the whole was overwhelmingly positive. Across all five Hackathons, and all participating teams, the Hackathon programme was viewed as an empowering experience, providing what was seen as a unique opportunity for people with cognitive disabilities to get their voice heard, and to actively make a practical contribution to driving forward changes in the design and delivery of digital tools and services. Participants also viewed the design and implementation format of the programme in favourable terms. The introduction of a flexible format, with participation times spread over a relatively long period, rather than compressed into a short one- or two-day event, was particularly welcomed by participants whose everyday lives and circumstances are relatively unstructured and sometimes chaotic. The flexible format enabled them to join the Hackathons as and when it suited them. However, language was cited as an issue in terms of Hackathon delivery. Many of the participants did not have a second language. Whilst this was not a problem in terms of the opening and mid stages of the Hackathons, the final collaborative stage of the Hackathon – when teams are brought together to collaboratively review their work – presented challenges for those without English as a second language. Related to this, another aspect identified early on as a potential challenge was the 'language of presentation'. In the early stages of the programme some participants felt that the Hackathons were being presented in a too technical or academic language. However, this was addressed as the programme developed.

3.4 EVALUATION OF THE OPEN TOOLKIT

As noted above in Section 2.2 evaluation of the Open Toolkit is carried out mainly through Tasks 3.2 and 3.3. It was implemented through the scenarios' validation work carried out in the Co-Labs and through the five 'Hackathons' implemented in LIVE IT (as discussed above in Section 3.3). The evaluation methodology followed a similar approach to that used to evaluate the tools deployed in the Co-Lab scenarios validation experiments, i.e., structured walkthrough with the Toolkit; observational analysis of Toolkit use; user survey, including the item System Usability Scale and feedback focus groups and interviews. A total of 47 people with cognitive disabilities took part in the Open Toolkit evaluation through the Co-Labs scenarios validation activities, as well as an additional 10 experts including IT developers. 200 participants in the Hackathons contributed to the evaluation of the Open Toolkit, including 25 experts from the IT and 'design for all' communities.

Overall, user feedback on the Open Toolkit was positive. Table 10 shows the mean scores for the Open Toolkit on the 'usability and usefulness' and 'emotional response' evaluation measures.

Table 10 - Overall rating of the Open Toolkit

Usability & Usefulness (scale 1-5)			**Emotional response (scale -3 - +3)**			
Meet my needs	Easy to use	Able to carry out task	Overwhelmed-In control	Distracted-Focused	Frustrated-encouraged	Stressed-relaxed
3.9	4.1	4.1	2.3	2.0	2.7	2.3

As Table 10 shows, the Toolkit was highly rated on the 'meet my needs', 'easy to use' and 'ability to carry out task' measures as well as on the four 'emotional response' measures. However, significant variations were recorded between the different Co-Labs, with participants in the UK Labs rating the Toolkit less positively than average. This in turn reflects the different profiles of participants in the Labs, the presenting digital and cognitive challenges faced and the context of use. Younger people were likely to score the Toolkit lower on usability and usefulness and emotional response, as were people who presented with learning issues, including numeracy and literacy.

Four key positive attributes of the Open Toolkit were highlighted by the evaluation:

- The design and functionality of the platform and tools
- The ease of navigation
- The range and usefulness of the resources provided in the Toolkit and their capacity to meet user needs
- The collaborative and participatory environment created by the Toolkit.

A key positive attribute of the tool highlighted in the evaluation was the simple design, with the elimination of unnecessary text and use of visual cues, as observations from participants show:

> *"The main benefit of the toolkit is its simple design. It makes it useful. It follows the basics of easy-to-read design methodology. The small amount of graphic and written information on the interface screen facilitates caregivers and people with cognitive difficulties to interact with it and understand how to use it and find something useful."*
>
> *"I like the website; it is easy to follow and clean."*
>
> *"It's nice, no colours and no distractions".*

This in turn was seen as contributing to help users to navigate through the menu and select useful tools:

> *"The design allows the end user to be directed quickly to the main menu and select the tool of his or her preference. The simple design allows less steps that could be easily memorized by our end users".*

The majority of users considered the resources provided in the Toolkit to be comprehensive and useful:

> *"The available resources look excellent."*
>
> *"a very comprehensive offering".*

Participants in the evaluation referred to the potential strategic importance of the Toolkit as potentially an effective way of bridging the current disconnections between the online world and the lived experience of people with cognitive disabilities. This was seen as particularly significant in light of the tool and Toolkit's potential to reduce the bureaucratic obstacles to engaging in digital life. As one participant observed:

> "(the Toolkit can enable) *utilization in public services for bureaucratic work like completing forms or reading online written instructions in order to facilitate those tasks for people with cognitive difficulty. Even in cases that they could not fill the form or do paperwork on their own the toolkit educates them through those processes and its very important for their social inclusion, because it gives them a taste of how the world formulates its functions in the social services' sector".*

Also highly valued were the functionalities and tools incorporated in the Toolkit to support user feedback and co-creation of content:

> *"It's so valuable to have a collated list of tech. The rating system is very useful and the categorization of the tools by skill set is also very useful."*
>
> *"Having reviews from people who are also looking for it for dyslexia help is really helpful".*

The evaluation also highlighted a range of issues and challenges for Toolkit users, together with suggested recommendations for improvement. A key challenge identified was the extent to which the advisor tool and the Open Toolkit have the flexibility to cater for the wide diversity of potential user profiles and needs. A good example of this is the situation of people with cognitive disabilities who present with particularly severe issues, like severe neurological disorders. In such cases, direct engagement with technical platforms and tools is pretty much impossible and requires mediation between end user and technology. These 'intermediaries', for example caregivers, often have limited digital competences. The implications of this for toolkit design are, firstly, that the Toolkit should not be developed in isolation. To maximise its value there needs to be complementary support to enable intermediaries with low digital competences to learn how to use the Toolkit effectively. This is a 'structural' challenge that goes beyond the functionality of the advisor tool and Toolkit and poses questions about the extent to which the Open Toolkit should incorporate additional content on training and learning – for example including within the 'Catalogue' links to relevant digital skills training programmes. Secondly it implies more attention needs to be given to improving the 'user friendliness' of the advisor tool and toolkit from the perspective of intermediaries as well as from the perspective of people with cognitive disabilities. This is an 'operational' challenge that speaks to the need to introduce design features like improved visual cues, reduced use of text in instructions, clearer signposting and easy-to-use buttons. As two participants observed:

> *"The platform is good, but you need to enrol the educators and caregivers on it. We have some beneficiaries with dementia it is difficult for them to learn even new basic staff like click on a button, because simply they have started to forget the very basics like who they are or how to help themselves in everyday life. This toolkit consists of tools that*

could help them bridge the generation gap between the digital age and their generation, but they need extra help to use it even though it's easy to use".

"The involvement and digital education-training of caregivers is essential because they are environmental mediators between their beneficiaries and the world (and in this case the digital world). They are the people that people with cognitive problems trust the most, so they are not going to that digital unexplored for them world without their company, who is their caregiver. Caregivers and health care professionals are the keypersons who will enrol the targeted end users".

The importance of appropriate layout and interfaces was also highlighted from the Toolkit evaluation. This is also related to the theme of language, as discussed above in Section 3.3 on the evaluation of the LIVE IT Hackathons. Many participants in the Co-Labs and Hackathons took the view that the default position of tool developers is that English is the 'lingua technica'. As a result, not only the content of most assistive tools is developed in English, but also their interfaces. However, people with cognitive disabilities – particularly those with learning and reading issues – face additional challenges when using tools in languages other than their native language.

Other key areas for improvement identified by the Toolkit evaluation included provision of audio-visual and video-related support to enable users to 'read less, write less, interpret less, presume less', and a revision of the Toolkit rating system. Specific feedback was to change the rating system from a stars' system to an emoji system.

A total of 76 specific recommendations for improvements to the Toolkit were logged through the evaluation. These are summarised in Table 11. A majority of these suggested recommendations have been implemented in the revised version of the Toolkit.

Table 11 - Recommended improvements to the Open Toolkit

Problem	**Potential solutions**
Appearance	
The LIVE IT Logo is tiny	Increase size
No info on what the toolkit is for or how to use it	Introductory information in text, video and audio format, human or AI support
Font size is too small, both heading and body of text	Make adjustable by user, have in large size as standard
Font style	Make adjustable by user, have in Open Dyslexic font as standard

Font color	Make adjustable by user
Background is white	Make adjustable by user
Lack of visuals	Widgets, images, sign language
Buttons/links are unclear	Underline text buttons to make clear they are a link
Images/icons are too small	Make larger or adjustable
Contrast	Increase of the contrast (use of frames with darker color), so the functionalities are displayed distinctly.
Layout	
Layout doesn't make clear the best way to navigate	A more considered design will enable users to make a journey through the toolkit. Navigational structure needs to be user friendly with intuitive icons and arrow to direct users. Voice navigation could be added for those who have difficulty understanding even this very basic menu design.
Layout is not interesting/appealing	Consider interactive and multimodal content
Things are too close together	Have more space between tools/content
Toolkit selection and type of tool should be next to each other (side by side)	Tools navigation
The cookies bar should be moved in the lower part of the page, because there are functionalities hidden	the cookies bar does not seem to be responsive. In smaller screens the text should be wrapped in the next line and the Accept Button should not be moved
Language	
Lack of language options	Easy to switch content to different languages
Too much text/information - dense text	Option for summary or detailed text, supporting images. Easy to read version of the toolkit – break text up with images/videos
Language is too complex	Option for summary or detailed text, supporting images. Easy to read version of the toolkit

No option for audio descriptions	Audio options in different languages and accents
Specific Advisor Tool Feedback	
Drop down menu is not user friendly	Consider alternative ways of filtering, for example, multiple choice buttons
No info on what each type of tools is/means	Option to click on heading for more details (audio or text based). Small image/logo of each tool that represents in short, what it does
Limited filtering options	Enable users to filter the results by a variety of grouping options, for example cost, language, free/paid, user need, rating. One suggestion is as follows: The first thing users would do is answer a series of simple questions to establish what their needs are. These would have multiple choice answers. Users would have the choice to read and/or hear the questions. Ideally there would be a chatbot doing this. Or even better an interactive virtual receptionist (avatar)* The questions would enable the list of tools to be narrowed down to those that meet the user's needs. Example questions: What device / browser are you using, what OS? What challenge / problem do you need to solve? difficulty reading difficulty with speech and grammar difficulty with time management Problems with numbers etc. Free or paid tools? Once you've answered the questions you enter your virtual library. Here you have access to a range of tools which you can test without the need to download / install / buy them.

	There is a chat room where you can communicate with other users. The chatbot (Helpdesk) is available to offer help and support, including access to tutorials.
Unclear information on each tool	If tools were available and preloaded in the library even better (and safer)
Unclear which tools are free or paid	Have this indicated next to each tool
What does rate mean?	Choose a word more easily understood
Rate button is red	Choose a color suitable for the colorblind
Focus button	"Since we aim to provide the user with a tool, the most focused icon, button or link should be the one that leads to the tools itself"
Icons (app logos) are too small	make them bigger/adjustable
Icons (app logos) are not links to info	insert link to app page
Tools should be ordered by rating (best to worst)	Tools' order
Star rating is unclear and too small	Larger emoji rating tool
How do we know why people rated the way they did?	Users may find interesting to add comments as well. You may add another tab named "Add comment (optional)"
There is no way to prevent a user from rating again.	Users need a login?
After rating it is not efficient to immediately change the position based on the rating of the tools because in this case a user can re-rate the same tool.	Rating presentation

“Top 5” could be the default value, and if user wants to see all tools they can choose “All”. Consider showing the “Top 5” in a different section.	Top 5 as default
In the tab all, you could add a scroll bar.	Scroll bar
You can only see the results in a list	Add grid view option
"I did wonder how the top 5 lists were curated so it would be good to know what criteria was used to determine top 5."	Top 5 explanation
Does 'disability' need to be repeated in each tool description?	Decrease use of word disability
It's unclear how to submit tools	Submit tools clarification
Tools navigation	"...when you click on the name of the tool it should bring you to more information and then you've the link there to go onto the website. It would be less buttons and an easier flow."
The toolkit is complicated to use	The Toolkit should open in a new window. The app ‘more’ button too could open in a new window.
Tools require download and installation	-The advisor tool could be an online space to spend time trying tools, sharing findings and ideas with others and getting support and guidance
Arrangement	-Could be arranged according to relevance to issues, browsers, OS, platform or Access Tech (Android, iPhone, Apple iOS, Windows, Chrome, Edge, Firefox and so on)
Number results	Number the results

Empty sidebar	Remove empty sidebar or add functionalities in it
Capabilities	
No support	Live chat with help
No clear way to communicate with other participants	Contact other participants/users
No way to communicate with experts	Contact special educators or experts, input from experts
No clear or easy way to share tools	An easy copy and paste button/function
Videos	
Do they have subtitles?	Subtitles
Are they clear and concise?	Clarity
Are there enough?	Include more videos with captions, especially to show users how to interact with the toolkit
Technical issues	
On every load/refresh the sidebar is slowly transitioning	Consider removing hamburger trigger
Empty sidebar	Remove empty sidebar or add functionalities to it
Consistency	Create consistency and use common UI elements
Site doesn't meet Web Content Accessibility Guidelines	Check against Web Content Accessibility Guidelines
More button	"When I clicked more and then returned to the tools list, it didn't go back to the category I was browsing, and I had to reload each time.

When you choose a category and press Go Back Button it returns you to the start and you have to re-choose the category.	You could rename the button to “Back to main screen” and add another one for “Back to results”
Consistency	Create consistency and use common UI elements
HTML5 semantic elements	Use HTML5 Semantic elements for proper code structure.
ARIA use	Use Accessible Rich Internet Applications (ARIA) set of attributes to improve accessibility for people with disabilities
When rating is done it is not useful to throw you to “More”	After rating
"The 3 lines beside the logo at the top are do not really do much. For me, they just make the logo smaller which is not necessary."	Remove them
The Notification (bell) icon does not seem to be working.	It should be either hidden when there are no notifications, or the user should be redirect to a page with past notifications.
What’s missing?	
No link to social media channels	Links to YouTube, Instagram, FB, Linked in need to be added
"I think each section should also show people how to use built in accessibility tools on their devices."	Built-in accessibility features

"One thing I thought was missing was showing people how to access the features already on their devices. Until I did an assistive technology course a few years ago, we weren't using our iPads to full potential at all as regards accessibility."	Access features guide
There should be a loading spinner added, so the user is acknowledged that the search is being processed.	Loading spinner
Browser extension	Make the Toolkit an extension for browsers, with all the features available in the same place.
Safety	
VPN	Free VPN could help due to concerns about hacking vulnerability cyber security
Installing app	Installing apps from a trusted source (in the virtual workspace with pre-loaded apps as a starting point perhaps)?
Virtual space	Virtual safe trusted space
Makerspace	
No link to YouTube	Add link to YouTube

3.5 ECONOMIC ANALYSIS

As noted in D5.3 – Sustainability Plan - our original intention was to use Social Return on Investment to assess the economic potential of LIVE IT and the Open Toolkit. SROI measures social, environmental and economic outcomes and uses monetary values to represent them[ix]. However, SROI requires financial proxies - indirect indicators that approximate for a direct indicator for which data is difficult to obtain. Proxies are normally obtained either through established databases – e.g., national data on unit health costs for an individual presenting with cognitive disabilities – or through primary data collection – for example a survey of people with cognitive disabilities. Since neither proxy sources were available for LIVE IT we used 'Cost Consequence Analysis' (CCA) as an alternative to SROI. The big difference between CCA and SROI is that CCA "does not attempt to summarise outcomes in a single measure (such as the

quality-adjusted life year) or in financial terms. Instead, outcomes are shown in their natural units (some of which may be monetary) and it is left to decision-makers to determine whether, overall, the treatment is worth carrying out"[x] CCA is generally used to compare two scenarios – the current status quo and an alternative scenario represented by the intervention for two key areas – User benefits and Business Impact. To do this for LIVE IT we reviewed and integrated evidence from desk research and from the evaluation data derived from LIVE IT's Lab activities and Hackathons. The results are summarised in the Table below.

Table 12 - LIVE IT CCA comparing 'implementation' and 'do-nothing' scenarios

Impact Area: User benefits	
Implementation scenario	**Do Nothing scenario**
Evaluation data from the LIVE IT Lab experiments shows that the use of tools to support increased web accessibility for PwCD has potentially significant user benefits. As an example, the responses of 50 PwCD to 26 different tools (e.g., Text to speech, Natural Reader, Speechify) across five LIVE IT Labs were collated and compared for usability and usefulness and emotional response. The mean score (on a maximum 100) for the tools combined was 76 on 'ability to meet my needs', 72 on 'ease of use' and 75 on 'ability to carry out tasks'. Low scores (below 50) were also recorded on the 'emotional response' indicators measuring stress, lack of control, distraction and frustration. These positive findings were reinforced through observational analysis carried out in the Labs. Scaling these benefits up and out would likely make a significant contribution to improvements in mental well-being for people with cognitive disabilities. It would likely lead to improved access to digital tools and services, increased quality of life and life opportunities and reduced social isolation. Ultimately this contributed to improved social inclusion	Current estimate is 142 million Europeans with a cognitive disability – dementia, dyslexia, ADHD, autism, learning and comprehension -19% of the population (Sources: WHO, European Parliament, Eurostat, PISA). 54% estimated to be digitally excluded. That's 77 million people – over 10% of the total population (Sources: Eurostat; EU SILC) This represents an enormous challenge to EU digital inclusion goals. The EU's digital targets for 2030 ('2030 Digital Compass') set a minimum of 80% of the population with basic digital skills, with ambitious targets for digitisation of public services – 100% access to e-health records; 80% citizens using digital ID. The underlying 'digital principles' that support these targets include 'freedom of choice' (a fair online environment) 'solidarity and inclusion (universal access to internet, to digital skills, to digital public services and to fair working conditions) and 'participation' (all citizens engage in the democratic process at all levels). None of these principles will be fully operationalised and targets met if a significant segment of the population – PwCD - remains digitally excluded.
Impact Area: Business Impact	
Implementation scenario	**Do Nothing scenario**
The LIVE IT Lab activities and experiments, supported by evaluation of the 'Hackathon' series suggests that participating PwCD increased basic digital skills, expanded access to online tools and	If no additional EU-level action is taken to support the digital transformation, the cost is estimated at €315 billion in 2021, and would be expected to grow over time to reach €1.3 trillion

services and had benefits for participants in terms of increased sense of self-efficacy and locus of control. Although difficult to quantify there is strong evidence to suggest these benefits have a significant economic multiplier. As an example, a 2018 Report by the Good Things Foundation, UK suggests the economic benefits over a ten-year period for the UK are: Time savings through online transactions - £0.11b per annum Earnings benefits from increased skills - £57m p.a. Employability benefits - £31m p.a. Transaction benefits - e.g., online shopping - £0.11b p.a. Health - e.g., online booking £14m p.a. Communication benefits - £40m p.a. Digital efficiency - £48m p.a. Corporate - e.g., reduction in skills shortages £0.15b p.a.	by 2033. (Source: European Parliament Thinktank, 2022). The economic impact of mental health problems – including cognitive issues - is estimated at more than €600 billion across the EU. This is accounted for by spending on health care, about 1.3% GDP or €190 billion, and social security programmes (1.2% or €170 billion euros. But the most important economic impact is due to lower employment and productivity of people with psychological disorders – which accounts for 1.6% GDP or €260 billion. (Source: OECD, 2021)

As Table 12 shows Cost Consequence Analysis suggests that the potential economic benefits associated with increasing web accessibility for people with cognitive disabilities are considerable. Considering estimates of the economic impact on the EU of mental health problems – 600b euro in 2021 – a rough estimate of the economic costs of exclusion for people with cognitive disabilities works out at around 15,700 euro per capita. These costs are mainly associated with spending on health care and social security as well as lower employment and productivity of people with cognitive disorders. They are likely to be considerably reduced as a result of increased web accessibility since much of the economy is now driven by online services and transactions. In addition, there are considerable direct economic benefits associated with increasing online opportunities for people with cognitive disabilities. As the data from the UK show, these are mainly associated with transaction benefits, communication benefits and skills improvements. Against this background, LIVE IT provided positive outcomes for 355 participants in the Co-Labs and Hackathons, including 'measurable' improvements like increased digital and media literacy skills, as well as 'intangibles' like increased self-esteem. Taking the absolute costs of LIVE IT into consideration – a total of 571,530 euro, the per capita cost of participation works out at 1,609 euro per participant. This suggests a potential rate of return for LIVE IT of 9.75. Of course, this is a very crude estimate, but it does suggest that LIVE IT has provided value for money.

4 THE BIG PICTURE: LIVE IT OVERALL SUMMATIVE EVALUATION

This concluding section draws together the evaluation results as presented in the preceding sections to provide an overview summative evaluation of LIVE IT. It begins by applying Theory of Change analysis to assess how far LIVE IT has travelled on its expected 'change journey' and to what extent the project's expected outputs, outcomes and impacts have been achieved. This is followed by a summary of the key messages from the evaluation and their implications.

4.1 THEORY OF CHANGE ANALYSIS

As discussed above in Section 2.3 LIVE IT's Theory of Change tells the project 'story' – from the 'presenting problem' it addresses through to the change it hopes to make on that problem at the end of the project and beyond (i.e., the project's expected 'impacts'). The Theory of Change specifies the critical success factors (activities and outputs) needed to deliver the project objectives; the immediate outcomes (changes in awareness, attitudes, knowledge and skills) expected to be realised as a result of the project - for example improved understandings of what works for the digital inclusion of people with cognitive disabilities; the intermediate outcomes (changes in individual behaviours and in the structures and processes of organisations) expected to be realised - for example digital platform designers increasing the accessibility of the technologies they design - and the potential 'high level' longer term impacts LIVE IT is expected to contribute to, i.e.:

- Better design of online content and services
- Increased digital inclusion of PwCD.

An important part of Theory of Change (ToC) analysis is identifying the 'mechanisms' through which LIVE IT delivers its outcomes. Mechanisms can be defined as:

> 'underlying entities, processes, or structures which operate in particular contexts to generate outcomes of interest' (Astbury and Leeuw, 2010)[xi]

LIVE IT is intended to encourage its target groups – people with cognitive disabilities, tool developers, policy-makers - to make and sustain different choices – for example choosing to participate in Hackathons; choosing a particular assistive tool to solve a digital challenge. Making these choices requires a change in the participant's 'reasoning' (for example the values, beliefs, attitudes, or the logic they apply to a particular situation). It also requires a change in the 'resources' participants have available to them. LIVE IT provides information, skills, material resources, and support which will in turn increase participants' individual resources (in information, skills etc.). This combination of 'reasoning and resources' is what enables LIVE IT to 'work' and is defined as a project 'mechanism'.

Three 'primary mechanisms' have been developed for the LIVE IT summative evaluation. These are summarised below in Tables 13 to 15. Each Table provides:

- A description of the mechanism
- The outcomes and impacts expected
- Evaluation evidence to support the mechanism
- Evaluation evidence to refute the mechanism.

Table 13 - Mechanism 1 - Knowledge creation, diffusion and mind-set change

Mechanism 1: Knowledge creation, diffusion and mind-set change	
Description	LIVE IT completes baseline research on state of the art in tools to support online accessibility for people with cognitive disabilities. Supported by in-depth lifeworld analysis that captures and documents digital life from the lived experience of people with cognitive disabilities, this leads to a significant improvement in the knowledge base on the digital challenges PwCD face, their digital needs and what is likely to work to increase digital inclusion. This improved knowledge base feeds into the development of design for all scenarios that are subsequently explored in the LIVE IT Co-Labs through experimentation. LIVE IT's supporting Hackathons and Makerspace further deepen and broaden the evidence base on what works. The outputs from Co-Labs, Hackathons and Makerspace feed into the development of the LIVE IT Open Toolkit. Utilisation of the Toolkit by stakeholders leads to an increase in know-how and expertise, which in turn contributed to changing perceptions about how web accessibility can be delivered for people with cognitive disabilities. LIVE ITs dissemination activities contribute to wide dissemination of the evidence, knowledge and learning from the project. This further contributes to changing perceptions about how web accessibility can be delivered for people with cognitive disabilities.
Outcomes and Impacts	Increased awareness of needs and barriers to digital inclusion for PwCD Increased awareness of what works in supporting web accessibility for PwCD PwCD improve their understanding of what works to support their web accessibility Policymakers, IT designers and developers and organisations providing services to PwCD change their perceptions of how tools and services should be designed and developed to increase web accessibility
Supporting evidence	180 literature review items and good practice cases intensively reviewed 150 PwCD involved in lifeworld analysis 233 good practice cases of platforms, development tools and authoring systems assessed Learning from the review and lifeworld analysis leads to identification of over 84 challenges, 1178 tools and development of 72 validation

	scenarios in which 351 experiments to develop solutions to the challenges are carried out 355 stakeholders – mainly PwCD – involved in co-design activities in Labs; 200 in the Hackathons 125 different tools evaluated 200 proposals for solutions are developed to support improved web accessibility for PwCD Open Toolkit developed incorporating Advisor Tool, Catalogue of validated good practices, stakeholder directory and design and policymaking guidelines 250 stakeholders engaged in LIVE IT community 250 stakeholders participate in final conference 4 webinars and 4 local seminars delivered 5 leaflets/newsletters distributed Article on state-of-the-art review published in prestigious academic peer-reviewed journal
Refuting evidence	Social media and website traffic relatively low Low level of engagement in the LIVE IT Makerspace Limited involvement by key stakeholder constituencies – particularly the IT commercial sector and policy-makers – in LIVE IT activities

As Table 13 shows, Mechanism 1 - Knowledge creation, diffusion and mind-set change – focuses on the contribution made by LIVE IT to raising awareness about the digital challenges people with cognitive disabilities face, expanding the knowledge base on challenges and user needs, creating and diffusing new knowledge about what works in this field and as a result positively changing mind sets to support 'design for all' thinking in policy, practice and the commercial IT sector circles. There is considerable evidence from the evaluation to support the conclusion that LIVE IT achieved significant progress in this part of its 'change journey'. The project has clearly significantly expanded the knowledge base in the field of web accessibility. It engaged a large number of people with cognitive disabilities in co-design work in the LIVE IT Co-Labs and Hackathons.

This work led to a large number of proposals for improvements to accessibility tools, as well as feeding into the ongoing development of the Open Toolkit. The Toolkit is a major – probably unique – resource for a broad range of stakeholders, including people with cognitive disabilities, the organisations that represent them, the IT industry and policymakers. It distils and valorises all of the evidence and learning from the project.

What is less clear is the extent to which this evidence, this learning and these resources have contributed to changing stakeholder perceptions, attitudes, skills and mind-sets. On the one hand, the evaluation showed that people with cognitive disabilities increased their level of knowledge and understanding of the digital options available to them improve their digital life. User feedback on the Open Toolkit also supports the conclusion that, in particular, the LIVE IT Advisor Tool further contributes to enabling people with cognitive disabilities to recognise the

potential value of different types of accessibility tools. Set against this is evidence of the relatively low engagement of key constituencies – notably the IT sector and policy-makers – in LIVE IT activities. In addition, public engagement has been low, as evidenced by data on website traffic and social media.

Table 14 - Mechanism 2 - Strategic decision-making

Mechanism 2: Strategic decision-making	
Description	Policymakers, tool designers and developers and service providers find out about LIVE IT through the project website, dissemination events, partner awareness-raising actions and networks. They see that LIVE IT fills a gap in their needs and sign up for the platform and Knowledge Community. Using the Open Toolkit and participating in the Knowledge Community increases their understanding of what works in web accessibility and improves their competence in design for all, giving them the confidence to apply it in practice. After LIVE IT has finished, they apply their new competences in their strategic thinking and operational practice. At the same time, people with cognitive disabilities participating in the Co-Lab experiments, Hackathons and Makerspace acquire new understanding and knowledge about what works in web accessibility, as well as the competences to make strategic decisions that will improve their access to digital opportunities. They apply this in their everyday lives and use the Advisor Tool to support their decision-making. Over time, Open Toolkit users provide feedback on content and use the ratings system to add to the evidence base on what works. This further contributes to strategic decision-making
Outcomes and Impacts	PwCD use the Open Toolkit to make decisions about which tools to use to meet their needs More extensive use of evidence-based knowledge in digital inclusion decision-making, design and practice by organisations supporting PwCD, tool designers and policymakers
Supporting evidence	Evaluation data suggests that participating PwCD in the Co-Labs and Hackathons are already applying what they have learned by selecting specific evaluated tools to meet their accessibility needs Open Toolkit user feedback suggests many PwCD intend to continue using the Advisor Tool to make decisions on which accessibility tools to use in the future. Positive feedback on Open Toolkit from IT and design for all communities
Refuting evidence	Low level of engagement in Makerspace activities Low utilisation of user feedback and ratings tool in Advisor Tool Limited involvement by key stakeholder constituencies – particularly the IT commercial sector and policymakers – in LIVE IT activities No strong evidence that the Open Toolkit is changing behaviours

Mechanism 2 - Strategic decision-making – focuses on 'intermediate outcomes' **–** changes in individual behaviours and in the structures and processes of organisations that can be attributed to the actions of LIVE IT. It looks for evidence that changes in awareness, attitudes, knowledge and skills are then converted into actions, for example using the LIVE IT Advisor Tool on a routine basis to select and use tools that will increase web accessibility.

From the available evaluation data, there is insufficient evidence to support the conclusion that LIVE IT has contributed significantly to behaviour change. On the one hand, feedback data from people with cognitive disabilities who participated in the Co-Lab experiments, Hackathons and Makerspace suggests that some participants are using the Open Toolkit to help them select and use tools that will increase their web accessibility in areas that are important to them. There is also evidence that these users intend to use the Toolkit in the future. However, the evaluation data suggests that this group of users is in a minority. The low level of usage of the Toolkit feedback and ratings functionalities, coupled with the relatively low engagement of participants in the LIVE IT Community reinforces the conclusion that there is some way to go before LIVE IT makes the transition from 'knowledge provider' to 'behaviour change' catalyst. This conclusion is more strongly reinforced when it comes to other key stakeholder groups. Given the low level of engagement from key stakeholder communities from the commercial IT and policy sectors, it is unlikely that any significant changes in individual or organisation decision-making in these broader sectors can be attributed to the project at this time.

Table 15 - Mechanism 3 - Paradigm shift

Mechanism 3: Paradigm shift	
Description	Through its networking and dissemination activities, the LIVE IT Open Toolkit is shared across a wide range of key stakeholder communities. A large number of PwCD and their supporting organisations use the Toolkit to improve the digital accessibility of PwCD. This engages PwCD more fully in digital life, opens up more life opportunities, for example education, employment, entrepreneurship and services access, and increases their social inclusion. Policymakers, IT designers and developers and systems providers use the Toolkit to make radical changes in the way tools, systems and services are designed. As a result, LIVE IT contributes to a disruptive and transformational shift in IT and policy culture.
Outcomes and Impacts	Better design of online platforms, tools, content and services for people with cognitive disabilities across the board Increased digital and social inclusion of PwCD.
Supporting evidence	No concrete evidence of large-scale change CCA suggests significant future impacts Replication analysis supports scalability potential Sustainability plan sets out transferability strategy
Refuting evidence	Although LIVE IT significantly exceeded its target for participating PwCD engaged in co-creation work in the Labs, Hackathons and Makerspace, the number of participants is a small fraction of the

	number of people with cognitive disabilities estimated to be digitally excluded The LIVE IT Knowledge Community did not attract significant numbers of PwCD not participating in the project LIVE IT has not engaged with enough stakeholders, and in particular IT development and policy decision-makers, across a broad range of constituencies As a result, a critical mass for change has not been achieved

Mechanism 3 - Paradigm shift – focuses on the high level and longer-term impacts of LIVE IT. It looks for evidence that behavioural and organisational change in aggregate are contributing to systemic changes at the broader societal level – in particular whether there is any evidence that LIVE IT has contributed to the increased social and digital inclusion of people with cognitive disabilities. There is no evidence from the evaluation that this type of disruptive, transformational and systemic change can be attributed to the project at this stage. However, work carried out for the project sustainability plan shows significant potential impacts and potential scalability of LIVE IT.

Taking all three mechanisms into consideration, the key message from the Theory of Change analysis is that LIVE IT has progressed to the 'behaviour change' point of an optimal 'change journey'. Its contribution to the ultimate aim of supporting systemic change in the way digital tools and services are designed, in order to increase accessibility for people with cognitive disability, and thereby increase their digital and social inclusion, is at this stage an embryonic contribution. This conclusion is about what you would expect for a preparatory and pilot action like LIVE IT. Its purpose was to explore the relevance, potential and feasibility of future actions within this field of web accessibility for people with cognitive disabilities. Given its short time scale and limited resources, it was never going to promote dramatic changes in policy or practice. What it has done is to justify the case for further actions in this field; provide evidence on what works; develop extensive resources to help future researchers, policymakers, IT developers and service providers change the way tools and services are designed and laid the foundations for the dissemination and application of the learning derived from its activities.

4.2 KEY EVALUATION MESSAGES AND IMPLICATIONS

- Analysis of the project outputs shows that, across most evaluation measures, LIVE IT has exceeded its targets, in particular for baseline research, the number of participants working in the LIVE IT 'Co-Labs', the accessibility tools evaluated in Co-Lab experiments and the outputs produced in the Labs.
- LIVE IT's outreach and dissemination activities present a mixed picture. On the one hand, the project reached or exceeded its targets in terms of webinars and local seminars delivered, conference presentations and peer-reviewed journal articles produced and attendance at the final conference. However, LIVE IT had limited success in reaching its broader target audience of stakeholder communities. In particular the level of involvement in the Makerspace was very low.

- The preliminary work carried out through baseline research, followed by intensive work with participants in the Co-Labs led to the mapping and analysis of around 85 challenges that prevent people with cognitive disabilities fully benefiting from digital life. These challenges were mapped against 1200 tools with the potential to deliver solutions to these challenges.
- Subsequent co-creation work carried out in the Co-Labs, involving 355 people with cognitive disabilities as co-designers and co-developers exploring the efficacy of a range of these tools, covered some 351 experiments in 72 different digital challenge scenarios. 125 tools to improve web accessibility for people with cognitive disabilities were systematically evaluated in the Co-Labs, on the basis of usability and usefulness and 'emotional response'. The results of this work fed into the development of the LIVE IT 'Open Toolkit'.
- In parallel, five Hackathons were implemented over the project lifecycle. 200 people participated in the Hackathons, in 38 different teams. These teams carried out further experiments with the suite of assistive tools, and with the LIVE IT 'Open Toolkit'. As a result of the Co-Lab and Hackathon work and, to a lesser extent the work done in the LIVE IT Makerspace, around 200 'interventions' were co-produced involving recommendations for changes to assistive tools, ideas for new tools and improvements to the Open Toolkit.
- The Open Toolkit is a major resource for a broad range of stakeholders, including people with cognitive disabilities, the organisations that represent them, the IT industry and policymakers. It distils and valorises all of the evidence and learning from the project and provides a range of tools to support web accessibility, including an 'Advisor Tool' to assist users in making evidence-based choices on which assistive tools to use. Evaluation of the Open Toolkit was very positive.

- An economic evaluation of LIVE IT, using a 'cost consequence analysis' (CAA) methodology suggests that the potential economic benefits associated with increasing web accessibility for people with cognitive disabilities are considerable, going forward.

- Theory of change analysis shows that LIVE IT has progressed to the 'behaviour change' point of a 'change journey' that, ultimately, could lead to systemic, disruptive and transformational change in the way digital tools and services are designed. This position is what might be expected for a preparatory and pilot action like LIVE IT. The evaluation shows that the project has provided a justification for further actions in this field; developed a significant evidence base on what works; developed extensive resources to help future researchers, policymakers, IT developers and service providers change the way tools and services are designed and laid the foundations for the dissemination and application of the learning derived from its activities.

ANNEX I: Detailed summary of LIVE IT Co-Labs scenarios validation

Help and Support Lab, Thessaloniki

1) Find and use contact information	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
1a) Search for fire department phone number	10	8	3
1b) Search for police department phone number	10	8	3
1c) Search for ambulance phone number	10	8	3
1d) Search for correct contact information for a service or business	10	5	7
1e) Search for relative's phone number and send a text or a voice message about an emergency	10	5	7
2) Find an overnight pharmacy	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
2a) Search for clinics and pharmacies	10	6	8
2b) Search for related information/services	0	0	0
3) Navigate with Google Maps	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
3a) Use Google Maps to define a location	10	0	10
3b) Use Google Maps to plan a route	10	0	10

4) Obtain information via email exchange	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
4a) Email question or complaint to a business/service	10	0	10
4b) Email a reply	10	0	10

Revised scenarios

1) Obtain information via email exchange	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
1a) Send message via email and social media	5	5	0

New Scenarios

1) Find information online	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
1a) Search for information about favourite musician	1	1	0
1b) Search for information on security issues	0	0	0
1c) Search for information to plan a vacation	5	5	0
2) Find COVID-19 information online and share with others	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
2a) Write an email about COVID-19 protective measures	5	5	0
2b) Create a PowerPoint presentation about COVID-19 protective measures	5	5	4

3) Organise and plan using online tools	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
3a) Transfer homework online and use digital tools to organise	1	1	0
3b) Make a plan for the week using digital tools	5	5	3

UK Co-Labs, Bristol and London

Initial Scenarios

1) Finding help in an emergency	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
1a) Use video-based tools to find help	2	2	2
1b) Use image-based tools to find help	3	2	3
1c) Use T2S tools to find help	12	6	6
1d) Use S2T tools (on forms) to find help	12	6	6
1e) Use S2T tools (messages) to find help	4	2	2
1f) Use apps to find help	2	2	0
1g) Connect with others to find help and support	3	3	0
2) Preparing for an emergency	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
2a) Use video-based tools to prepare	2	2	0
2b) Use image-based tools to prepare	0	0	0
2c) Use T2S tools to prepare	4	2	2
2d) Create personal memos to prepare	1	1	0
2e) Use notetaking and mind-mapping tools to prepare	3	1	2
3) What to do in an emergency	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
3a) Use video-based tools to learn what to do	5	5	4
3b) Use image-based tools to learn what to do	3	4	3

3c) Use T2S tools to learn what to do	5	5	2
4) Optimise phone for use	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
4a) Learn how to do user-friendly device set-up	4	4	2

Revised scenarios

1) Individualised	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
1a) Participant-directed	7	7	7

New scenarios

1) Managing healthcare	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
1a) Managing healthcare via apps	4	2	2
1b) Managing mental health and well-being	4	2	2

Ireland, NUIG

Initial scenarios

1) Reading	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
1a) Use T2S tools (participant directed)	7	5	0
1b) Use T2S tools with assigned readings	7	5	3
2) Writing (via dictation)	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
2a) Use dictation/S2T to write an email	0	0	0
2b) Use dictation/S2T to write an essay (or other written assignment)	2	2	0
3) Writing (via typing)	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
3a) Write an email	11	11	2
3b) Write an essay (or other written assignment)	11	11	3

Revised scenarios

1) Reading & Writing (S2T)	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
1a) Use voice memos and S2T to generate transcript of voice memo (lecture)	1	1	0
1b) Use voice memos and S2T to generate transcript of voice memo (note-to-self)	1	1	0
2) Reading & Writing (read-aloud feature)	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
2a) Review one's own writing using read-aloud	3	3	0

2b) Use read-aloud with assigned readings	1	1	0

New scenarios

1) Managing healthcare	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
1a) Managing healthcare via apps	4	2	2
1b) Managing mental health and well-being	4	2	2

Portugal – Lisbon and Miranda do Corvo

Initial scenarios

1) Text-to-speech	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
1a) Use T2S with email	10	10	5
1b) Use T2S with news websites	10	10	2
1c) Use T2S with social media feed	10	10	1
1d) Use T2S with online train schedules	5	4	0
2) Translation	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
2a) Use translation tools with email	10	10	3
2b) Use translation tools with news websites	10	10	2
2c) Use translation tools with social media feed	10	10	0
3) Complex Tasks	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
3a) Book a COVID-19 vaccination appointment	5	3	0
3b) Use 2-factor authentication	5	3	0
3c) Buy items online	2	0	0

Revised Scenarios

1) Complex Tasks	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
1a) Search for information online	10	10	5
1b) Record voice memos in messaging apps	3	3	0

New scenarios

1) Understanding and interpreting numerical, text, and graphics information	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
1a) Use a national COVID-19 statistics website	1	1	0
2) Manage dates and appointments	Number of times used in co-lab	Number of successful implementations	Number of implementations with pivots required
2a) Use a digital calendar to save appointments	5	5	2

ANNEX II: Detailed map of digital challenges against tools

Filename: *tools mapped to challenges.pdf*

5 REFERENCES

[i] Pawson R, Greenhalgh T, Harvey G, Walshe K. (2005), Realist review--a new method of systematic review designed for complex policy interventions. J Health Serv Res Policy. 2005 Jul;10 Suppl 1:21-34.

[ii] Befani, B. (2012) 'Models of Causality and Causal Inference', in E. Stern, N. Stame, J. Mayne, K. Forss, R. Davies and B. Befani (eds), Broadening the Range of Designs and Methods for Impact Evaluations, DFID Working Paper 38, London: Department for International Development

[iii] https://www.w3.org/TR/2021/NOTE-coga-usable-20210429/

[iv] Gobble, M (2014) Design Thinking, Research-Technology Management, 57:3, 59-62

[v] IDEO - https://www.ideo.com/post/design-kit

[vi] Mayne ,J. (2012) "Contribution analysis: Coming of age?" *Evaluation*, 18(3), pp.270–280

[vii] *Brooke, J. (1986).* "SUS: a "quick and dirty" usability scale". *In P. W. Jordan; B. Thomas; B. A. Weerdmeester; A. L. McClelland (eds.). Usability Evaluation in Industry. London: Taylor and Francis.*

[viii] Klaar, Margus J (2014) How to Have Your Cake and Eat it too. An Introduction to Service Design, Bis Publishers, Amsterdam,

[ix] Source: 'A Guide to Social Return on Investment'. SROI Network, 2012

[x] https://www.nice.org.uk/Glossary?letter=C

[xi] Astbury B and Leeuw F (2010) Unpacking black boxes: mechanisms and theory building in evaluation. American Journal of Evaluation 31(3): 363–81.