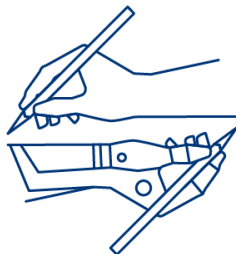




Funded by  
the European Union

Project: 101094364 — ITHACA — HORIZON-CL2-2022-DEMOCRACY-01  
EUROPEAN RESEARCH EXECUTIVE AGENCY (REA)  
REA.C – Future Society  
C.1 – Inclusive Society



**ITHACA**

AI To Enhance Civic Participation

**ITHACA**

**artificial Intelligence To enHance Civic pArticipation**

### D4.1: Platform test and evaluation plan

#### Work Package 4: Pilots' Implementation and Evaluation

<b>Authors:</b>	<b>CERTH</b>
<b>Status:</b>	<b>Final</b>
<b>Due Date:</b>	<b>31/12/24</b>
<b>Version:</b>	<b>Final</b>
<b>Submission Date:</b>	<b>20/12/24</b>
<b>Dissemination Level:</b>	<b>PU</b>

**Disclaimer:**

This document is issued within the frame and for the purpose of the ITHACA project. This project has received funding from the European Union's Horizon Europe Framework Programme under Grant Agreement No. 101094364. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the European Commission.

This document and its content are the property of the ITHACA Consortium. All rights relevant to this document are determined by the applicable laws. Access to this document does not grant any right or license on the document or its contents. This document or its contents are not to be used or treated in any manner inconsistent with the rights or interests of the ITHACA Consortium or the Partners detriment and are not to be disclosed externally without prior written consent from the ITHACA Partners. Each ITHACA Partner may use this document in conformity with the ITHACA Consortium Grant Agreement provisions.

(\*) Dissemination level. - Public — fully open (automatically posted online)

Sensitive — limited under the conditions of the Grant Agreement

EU classified — RESTREINT-UE/EU-RESTRICTED, CONFIDENTIEL-UE/EU-CONFIDENTIAL, SECRET-UE/EU-SECRET under Decision 2015/444

## ITHACA Project Profile

**Grant Agreement No.:** 101094364

<b>Acronym:</b>	<b>ITHACA</b>
<b>Title:</b>	<b>artificial Intelligence To enHance Civic pArticipation</b>
<b>URL:</b>	<b><a href="https://www.ithaca-project.eu/">https://www.ithaca-project.eu/</a></b>
<b>Start Date:</b>	<b>01/01/2023</b>
<b>Duration:</b>	<b>36 months</b>

### Partners

Short Name	Legal Name	Country
KT	KONNEKT ABLE TECHNOLOGIES LIMITED	IE
CERTH	ETHNIKO KENTRO EREVNAS KAI TECHNOLOGIKIS ANAPTYXIS	EL
UPAT	PANEPISTIMIO PATRON	EL
RtF	RAISING THE FLOOR	BE
SnP	STAMADIANOS KAI SYNETAIROI DIKIGORIKI ETAIREIA	EL
UniGraz	UNIVERSITAET GRAZ	AT
MNLT	MNLT INNOVATIONS IKE	EL
SIMAVI	SOFTWARE IMAGINATION & VISION SRL	RO
PEDAL	PEDAL CONSULTING SRO	SK
BMA	AGENTIA METROPOLITANA PENTRU DEZVOLTARE DURABILA BRASOV ASOCIATIA	RO
MARTIN	MESTO MARTIN	SK

**DOCUMENT HISTORY**

VERSION	DATE	CHANGES	RESPONSIBLE PARTNER
0.1	30/09/2024	Table of Contents	CERTH
0.2	22/05/2024	Complete version to be reviewed by WP4 partners	CERTH
0.3	02/12/2024	Version sent for internal review	CERTH
0.4	18/12/2024	Feedback incorporated	CERTH
Final	20/12/2024	Submission to EC	CERTH

## TABLE OF CONTENTS

<b>1.</b>	<b>Introduction to ITHACA</b>	<b>7</b>
1.1	Background of the ITHACA Project	7
1.2	Purpose of Deliverable D4.1	7
1.3	Structure of the Document	8
<b>2.</b>	<b>Framework for ITHACA Validation</b>	<b>10</b>
2.1	Definition of the Validation Framework	10
2.2	Key Performance Indicators (KPIs)	11
2.3	Testing scenarios	14
2.4	Pilot-Specific Adjustments	16
2.5	Additional Pilot-Specific Adaptations	18
2.6	Validation process for Pilot Sites	19
<b>3.</b>	<b>Implementation of the framework</b>	<b>20</b>
3.1	Evaluation stages	20
3.2	Time Plan for Pilot Activities	22
3.3	Recruitment and Engagement Strategies	23
3.4	User Support and Training Procedures	23
<b>4.</b>	<b>Evaluation Procedures for the Pilots</b>	<b>24</b>
4.1	Testing with experts (Phase 0): Cognitive walkthrough of platform wireframes	25
4.2	Controlled performance and usability testing (Phase 1)	27
4.3	Real-life performance and user experience testing (Phase 2)	30
<b>5.</b>	<b>Risk assessment and Mitigation</b>	<b>32</b>
5.1	Identified Risks for Pilot Implementation	32
5.2	Mitigation Strategies	33
<b>6.</b>	<b>Ethics and Data Privacy</b>	<b>34</b>
6.1	Ethics requirements and consent	34
6.2	Data privacy and GDPR compliance	35
<b>7.</b>	<b>Conclusion and Next Steps</b>	<b>35</b>
7.1	Summary of the Validation Framework	35
7.2	Recommendations for Pilot Implementation	36
7.3	Future Work and Follow-up Activities	36
	<b>References</b>	<b>37</b>
	<b>Annexes</b>	<b>38</b>
	Annex I: Pilot Site Details	38
	Annex II: Evaluation Forms and Questionnaires	44
	Annex III: List of Metrics	51

**Annex IV: Preliminary Testing Scenarios for Platform Functionalities..... 51**  
**Annex V: Template for Cognitive Walkthrough Test Plan for Expert Evaluation ..... 51**  
**Annex VI: Software Quality testing procedures..... 53**  
**Annex VII: User acceptance and experience testing (Phase 1 & 2)..... 62**  
**Annex VIII: Structure of the ITHACA Platform User Manual ..... 68**

**LIST OF TABLES**

Table 1: KPIs for Evaluating ITHACA Platform Stability, User Experience, and Ethical AI Compliance ..... 13

**ABBREVIATIONS**

<b>AI</b>	Artificial Intelligence
<b>GDPR</b>	General Data Protection Regulation
<b>HCI</b>	Human-Computer Interaction
<b>KPI</b>	Key Performance Indicator
<b>NGO</b>	Non-Governmental Organization
<b>STT</b>	Speech-to-Text
<b>TTS</b>	Text-to-Speech
<b>UAT</b>	User Acceptance Testing
<b>WP</b>	Work Package
<b>WCAG</b>	Web Content Accessibility Guidelines

## EXECUTIVE SUMMARY

This deliverable, **D4.1: Platform Test and Evaluation Plan**, presents a comprehensive framework for validating the ITHACA platform's functionality, usability, and impact through structured testing and pilot implementation. It outlines the methodology, tools, and criteria designed to ensure the platform aligns with its goals of enhancing civic participation and inclusivity.

The document begins by introducing the ITHACA project, funded under the Horizon Europe program, and its aim to leverage AI technologies to empower citizens in democratic processes. The testing framework builds on user and technical requirements defined in previous work packages (WP2 and WP3) and is tailored for two pilot sites: Brasov, Romania, and Martin, Slovakia. These sites were selected for their diverse socio-economic and technological contexts, providing a robust basis for evaluating the platform's adaptability.

Key sections of the document include:

1. **Validation Framework:** This section defines the metrics, tools, and methods to assess the platform's technical performance, usability, and accessibility. Emphasis is placed on iterative evaluation through two testing phases: controlled usability and performance testing (Phase 1) and real-life user experience testing (Phase 2).
2. **Pilot Implementation Plan:** A structured approach is outlined for deploying the platform, engaging stakeholders, and recruiting a diverse user base, including vulnerable groups. Strategies for user training, technical support, and community outreach ensure inclusivity and active participation.
3. **Evaluation Procedures:** The deliverable specifies the testing scenarios, KPIs, and success metrics for assessing the platform's impact. Methods include user surveys, focus groups, interviews with stakeholders, and system performance monitoring.
4. **Ethics and Data Privacy:** Ethical considerations, informed consent processes, and GDPR-compliant data handling protocols are described to safeguard user rights and ensure compliance with legal standards.
5. **Risk Assessment and Mitigation:** Potential risks, such as technical challenges and low user engagement, are identified, along with proactive mitigation strategies to ensure pilot continuity.

The deliverable concludes with recommendations for improving pilot implementation and outlines future activities, including final refinements and potential scaling of the platform to additional contexts.

The annexes provide supplemental resources and detailed information essential for understanding the methodologies, processes, and tools applied throughout the testing and evaluation of the ITHACA platform. They include **Annex I: Pilot Site Details**, **Annex II: Evaluation Forms and Questionnaires**, **Annex III: List of Metrics**, **Annex IV: Preliminary Testing Scenarios for Platform Functionalities**, **Annex V: Template for Cognitive Walkthrough Test Plan**, **Annex VI: Software Quality Testing Procedures**, **Annex VII: User Acceptance and Experience Testing Procedures**, and **Annex VIII: Structure of the ITHACA Platform User Manual**. These annexes ensure transparency and facilitate the replication of methodologies for similar projects.

This structured plan ensures that the ITHACA platform is thoroughly tested, user-centric, and aligned with its mission of fostering inclusive and effective civic participation.

# 1. Introduction to ITHACA

## 1.1 Background of the ITHACA Project

The ITHACA project, funded under the Horizon Europe programme, aims to enhance civic participation through the development and deployment of an AI-driven platform that empowers citizens to engage actively in democratic processes. The project's core objectives centre on creating a transparent, inclusive, and accessible platform for civic engagement, where citizens can voice their perspectives and collaborate with local governance entities in decision-making processes. ITHACA addresses contemporary challenges around democratic participation by implementing artificial intelligence (AI) solutions that not only facilitate dialogue but also ensure that civic engagement is secure, equitable, and respectful of individual privacy rights.

Work Package 4 (WP4) of ITHACA plays a pivotal role in achieving these goals by focusing on the implementation and evaluation of the platform through pilot deployments in two distinct locations: Brasov, Romania, and Martin, Slovakia. WP4 ensures that the ITHACA platform meets the technical and user requirements defined in earlier work packages (WP2 and WP3). By anchoring the project in real-world contexts, WP4 provides an essential framework for assessing the platform's usability, functionality, and overall impact on the communities it serves. The pilot activities within WP4 are designed to test the platform's robustness in diverse socio-economic environments, examining its adaptability to specific local needs and the effectiveness of its AI-based tools in promoting civic engagement.

## 1.2 Purpose of Deliverable D4.1

Deliverable D4.1, titled *Platform Test and Evaluation Plan*, establishes the framework for validating the ITHACA platform through structured testing and evaluation in both pilot sites. This deliverable is necessary for ensuring that the platform not only fulfils its intended technical functions but also aligns with the project's broader social objectives. Specifically, D4.1 provides the following contributions to the project:

- **Defines the validation criteria** by outlining the metrics, tools, and procedures needed to assess the platform's technical performance, user satisfaction, accessibility, and security, guiding the evaluation of the platform's functionality and ensuring it meets the needs of both general users and vulnerable users' communities.
- By setting the standards and evaluation methodologies, it **links to pilot implementation** by directly informing the deployment of the platform in Brasov and Martin, specifying validation approaches for each pilot and taking into account local demographics, technological infrastructure, and engagement requirements unique to each site.
- The evaluation framework **supports iterative improvement** through a three-phase approach to testing, with the first two phases (0 with expert evaluation and 1 with users in controlled environment) scheduled for Q1 2025 and focused on platform performance and short usability tests to identify immediate technical adjustments, and the second phase, following these adjustments, involving full deployment and real-life testing to enable a more comprehensive assessment of the platform's impact on civic engagement.
- **Aligns with WP4 activities** by providing the basis for structured intervention processes, recruitment strategies, and data collection methodologies, setting a clear timeline and

framework for the execution of pilot activities and ensuring that each phase of the project aligns with overall project timelines and objectives.

## 1.3 Structure of the Document

This document is structured to provide a comprehensive framework for the testing, deployment, and evaluation of the ITHACA platform across its two pilot sites. Each chapter has been designed to address specific aspects of the validation process, ensuring a systematic approach to assessing platform functionality, user experience, and overall impact.

- **Chapter 1: Introduction** outlines the background of the ITHACA project, with a particular focus on Work Package 4 (WP4) and the objectives of D4.1 within the larger project framework. This chapter contextualises the deliverable within ITHACA's goals, describing the rationale behind the platform's validation requirements (as described in D2.1) and linking these to pilot site deployment.
- **Chapter 2: Framework for ITHACA Validation** details the core validation framework that will guide platform assessment. This chapter defines key performance indicators (KPIs) and testing scenarios and explains how the validation framework will be applied and customised for each pilot site to accommodate site-specific characteristics, ensuring that the evaluation process aligns with both technical and user requirements.
- **Chapter 3: Pilot Implementation Framework** presents the process for implementing the platform at each pilot site. This chapter includes an intervention process outline, a time plan with key milestones and an overview of user support and training resources. It ensures a structured and consistent approach to pilot setup and deployment across both sites. The recruitment and engagement strategies for user participation are drawn from WP2 strategies already discussed in D2.1, the social context of the pilot sites presented in D1.3 and the best strategies as identified in D1.1.
- **Chapter 4: Evaluation Procedures for the Pilots** describes the methodologies for evaluating the platform's technical and functional performance. It includes the software quality requirements, component and integrated platform testing procedures, security and stress testing, and user acceptance testing (UAT). Success metrics for each aspect are provided to ensure clear benchmarks for performance and user satisfaction.
- **Chapter 5: Risk Assessment and Mitigation** identifies potential risks related to pilot implementation, including technical, logistical, and user-related risks. Mitigation strategies and contingency plans are discussed, providing a proactive approach to managing risks and ensuring pilot continuity.
- **Chapter 6: Ethics and Data Privacy** addresses ethical considerations and data privacy protocols. This chapter ensures that pilot activities adhere to ethical standards, focusing on participant consent, data handling, and compliance with GDPR and other relevant regulations to protect user privacy.
- **Chapter 7: Conclusion and Next Steps** summarises the key takeaways from the validation framework and offers recommendations for pilot implementation. Future work, follow-up activities, and potential platform adjustments are outlined to guide subsequent project phases and further developments.
- **The Annexes** in this deliverable serve as a foundational repository of detailed documentation, tools, and resources integral to the platform's testing and evaluation process. They include:

- **Pilot Site Details** (Annex I): Detailed profiles of the Brasov and Martin pilot sites, highlighting their unique socio-economic, technological, and demographic contexts. These profiles provide a basis for tailoring testing and validation procedures to each site's specific needs.
- **Evaluation Forms and Questionnaires** (Annex II): A collection of forms and instruments developed to gather user feedback on platform usability, accessibility, and functionality. These are designed to capture both qualitative and quantitative insights during the testing phases.
- **List of Metrics** (Annex III): Comprehensive metrics for evaluating the platform's performance, usability, and impact, including key performance indicators (KPIs) aligned with the project's goals.
- **Preliminary Testing Scenarios for Platform Functionalities** (Annex IV): Initial testing scenarios linked to anticipated platform functionalities. These scenarios are designed to validate core use cases and user interactions.
- **Template for Cognitive Walkthrough Test Plan** (Annex V): A structured template guiding expert evaluations of platform wireframes, focusing on usability, accessibility, and alignment with user needs.
- **Software Quality Testing Procedures** (Annex VI): Detailed protocols for assessing platform stability, security, and performance through rigorous testing methodologies, ensuring technical robustness.
- **User Acceptance and Experience Testing Procedures** (Annex VII): Guidelines for conducting comprehensive user acceptance testing and evaluating real-life user experiences to refine platform features and enhance usability.
- **Structure of the ITHACA Platform User Manual** (Annex VIII): An outline for the user manual that will support participants in navigating the platform's functionalities and accessibility tools effectively.

These annexes are integral to the document, offering in-depth insights into the methodologies and tools employed to ensure the ITHACA platform achieves its objectives of fostering inclusive, accessible, and impactful civic participation.

Each chapter builds upon the previous sections, ensuring a cohesive and logical flow that guides readers through the comprehensive validation plan for the ITHACA platform's deployment and evaluation at both pilot sites.

## 2. Framework for ITHACA Validation

### 2.1 Definition of the Validation Framework

The validation framework for the ITHACA platform is designed to systematically assess its technical performance, usability, and overall effectiveness in promoting civic participation. This framework establishes a comprehensive approach to ensure the platform meets both technical specifications and user-centric requirements, particularly focusing on fostering inclusive, accessible, and secure civic engagement.

At the core of this framework are metrics across several dimensions. Technical performance metrics, such as system stability, uptime, response times, error rates, and data processing accuracy, are critical for validating that the platform operates reliably under various loads and user demands. Usability metrics, focusing on ease of navigation, accessibility features, and user satisfaction levels, are especially important for ensuring the platform's accessibility to a diverse user base, including vulnerable groups. Engagement metrics measure user activity, participation rates, and content interaction levels, offering insights into the platform's effectiveness in fostering civic participation. Security and privacy metrics monitor data protection standards, GDPR compliance, and data handling practices, ensuring alignment with regulatory requirements and user expectations for data security.

To gather these metrics, the validation framework utilises a range of tools and data collection instruments. System monitoring tools track technical performance metrics such as server uptime, error logs, and load times, enabling real-time monitoring and rapid troubleshooting. User feedback mechanisms, including surveys, in-app feedback forms, and satisfaction ratings, facilitate direct input from users regarding platform usability and functionality, offering a window into user experiences and areas for improvement. Engagement analytics tools capture data on user interaction levels, participation trends, and usage frequency, helping assess the platform's success in meeting civic engagement objectives. In addition, security and privacy auditing tools help maintain GDPR compliance and privacy standards by tracking data access, managing consent, and safeguarding sensitive information.

The framework employs various methods to ensure a holistic evaluation of the platform. Testing procedures, including unit testing, component testing, and full integration testing, verify that each feature operates as expected. Stress testing assesses the platform's resilience under high-usage conditions. User Acceptance Testing (UAT), conducted with stakeholders and end-users, evaluates platform usability and satisfaction, providing valuable feedback on the user experience and identifying areas for refinement. The framework also adopts an iterative evaluation approach, implemented in two phases. Phase 1, scheduled for Q1 2025, focuses on initial performance and usability testing, which will help identify any immediate technical adjustments and refine user interfaces. Phase 2, spanning Q3 to Q4 2025, involves full deployment and real-life testing, allowing the project team to gather data on user engagement, civic participation, and the platform's impact on each community. In addition, qualitative feedback collection through interviews, focus groups, and open-ended survey questions provides in-depth insights into user experiences, perceived benefits, and challenges encountered with the platform.

This validation framework is carefully structured to align with both the user requirements defined in WP2 and technical specifications outlined in WP3, ensuring compliance with the project's dual objectives of technical reliability and user-centred design. Technical compliance is achieved through rigorous testing standards and continuous system monitoring, which ensure the platform's functionality, security, and efficiency by tracking metrics such as uptime, error rates, and data processing speed. User-centred design compliance is supported through usability and engagement metrics, along with qualitative feedback collection, to assess how effectively the platform meets user needs. These efforts ensure that the platform is intuitive, accessible, and encourages civic engagement among diverse community members, including those from vulnerable populations.

By integrating these components, the validation framework provides a solid foundation for monitoring, evaluating, and refining the ITHACA platform's performance in real-world settings.

## 2.2 Key Performance Indicators (KPIs)

The KPIs selected for the ITHACA platform's validation process are crafted to thoroughly assess its technical performance, user experience, and overall effectiveness in fostering civic engagement. These KPIs establish measurable targets across critical areas to ensure that the platform is reliable, accessible, and effective in encouraging community participation. Each KPI aligns closely with the project's objectives of technical reliability, user-centred design, and inclusivity.

In the realm of **Fairness and Ethics in AI**, ITHACA's emphasis on trustworthiness and ethical AI mandates metrics that evaluate fairness, accountability, and transparency. Fairness metrics, such as Predictive Parity and Statistical Parity, along with assessments of user trust and explainability, ensure that the AI functions equitably and that users have a clear understanding of AI-driven decisions. This focus on ethical AI supports a transparent civic engagement platform that respects user diversity.

**Privacy and Data Governance** is another essential area where data protection and compliance with GDPR are closely monitored. Metrics for this category include data privacy measures, anonymity levels, and success rates in privacy-preserving tests, all of which are vital for safeguarding user information. These privacy metrics help to build trust with users, ensuring that personal data is handled responsibly within the platform.

In line with ensuring transparency, **Explainability and Transparency** metrics measure user trust, explanation completeness, and the interactivity of AI explanations, helping users to understand and interact with AI systems. This supports the objective of creating a trustworthy platform where users can easily interpret AI-driven outputs, thereby enhancing transparency and trust in the technology.

**Robustness and Resilience** metrics focus on the platform's ability to withstand unexpected inputs and potential adversarial attacks. Metrics for robustness and error tolerance, along with measurements of resilience to security threats, are important for ensuring platform stability and security. These metrics support the platform's objective of delivering a dependable, secure environment, essential for user trust and continuous engagement.

The **Software Performance KPIs** are designed to evaluate the platform's technical stability, responsiveness, and reliability. Key metrics include system uptime, response time, and error rate. System uptime measures the percentage of time the platform is operational, ensuring consistent

user access. Response time tracks the platform's speed in addressing user requests, which contributes to a smooth and efficient user experience, ultimately enhancing satisfaction and engagement. Error rate assesses the frequency of system errors or crashes, a critical metric for maintaining user confidence. These KPIs ensure a robust technical foundation that supports reliable interactions, especially as the platform scales during real-world testing.

**User Satisfaction KPIs** focus on platform usability, accessibility, and overall appeal to participants. Usability score, based on user feedback, evaluates how easy it is for users to navigate the platform and complete tasks, fostering an inclusive experience that encourages engagement. Accessibility compliance measures adherence to standards like the Web Content Accessibility Guidelines (WCAG v.2.2) to ensure that users with disabilities can fully participate. User satisfaction ratings, collected through periodic surveys, provide an overall assessment of user experience, reflecting the platform's effectiveness in meeting user expectations. These KPIs highlight the importance of an inclusive and positive user experience, a core objective of ITHACA in promoting diverse community engagement.

The **System Reliability KPIs** measure the platform's capability to handle user demand and maintain performance across various conditions. Scalability performance evaluates the platform's capacity to accommodate increased traffic without performance degradation, which is essential for managing high civic activity. Data processing accuracy ensures that user inputs, feedback data, and engagement metrics are handled precisely, supporting reliable analytics and user trust. Security and privacy compliance metrics verify adherence to GDPR and other data protection standards, fostering user confidence by guaranteeing secure and confidential data handling. These reliability KPIs validate that the platform can maintain high performance, accuracy, and security as usage increases, supporting large-scale civic engagement.

**Alignment with Project Objectives** is achieved by structuring each KPI to reinforce ITHACA's commitment to a technically sound, user-focused platform that is adaptable to diverse community needs. Software performance KPIs ensure a reliable technical backbone, allowing for seamless user experiences. User satisfaction KPIs emphasise accessibility and user-centred design, which are essential for encouraging engagement across a broad demographic. Finally, system reliability KPIs guarantee that the platform can function at scale, protecting user data and privacy, thereby establishing ITHACA as a trusted tool for civic engagement.

The validation framework, built upon these KPIs, provides a structured and measurable approach to evaluate platform success. This framework ensures that the platform meets both functional and social objectives, fully supporting ITHACA's mission to advance democratic participation through inclusive, transparent, and reliable technology.

Table 1 reflects a comprehensive set of KPIs and metrics, aligning with technical performance, user experience, AI fairness, accessibility, engagement, and security/privacy objectives. Each metric supports a well-rounded validation framework, ensuring ITHACA's platform meets all functional, ethical, and social objectives.

**Table 1: KPIs for Evaluating ITHACA Platform Stability, User Experience, and Ethical AI Compliance**

Objective	Research Question	KPI	Metric	Target	Method of Data Collection
Ensure platform stability and reliability	Does the platform remain accessible and stable during normal and peak loads?	System Uptime	Percentage of uptime	≥ 99.5% uptime	System Monitoring Tools
		Response Time	Average response time	≤ 2 seconds per request	Performance Testing Tools
		Error Rate	Frequency of errors	≤ 3 critical, ≤ 10 minor errors/month	Error Logs
		Load Capacity	Concurrent user capacity	500+ users	Load Simulation Tools
Enhance AI-driven feature trustworthiness	Are AI-driven features operating safely and effectively?	AI Trustworthiness	Safety incident count	Zero safety incidents	AI Behaviour Monitoring
		AI Accuracy	Task completion success rate	≥ 95%	Benchmarking against expected outcomes
		AI Performance	Average response time	< 2 seconds	Response Time Tracking
		AI Accessibility	User satisfaction with AI	≥ 90% satisfaction for accessibility	User Feedback on AI
		AI Explainability	User understanding of AI decisions	≥ 80% understanding	User Surveys, Focus Groups
Improve user engagement and retention	How frequently and consistently are users interacting with the platform?	User Engagement	Monthly active users	≥ 75% of registered users	Platform Analytics
		Retention Rate	Continued user engagement	≥ 70% retention	User Activity Tracking
		Feature Adoption Rate	Usage of key features	≥ 70% of users adopt key features	Feature Usage Analytics
Provide a user-friendly experience	How easy is it for users to navigate and interact with the platform?	Usability Score	User feedback score	≥ 80% positive usability score	Usability Questionnaires
		Ease of Learning	User learning rate	≥ 85% rate platform as "easy to learn"	Post-use Surveys, Interviews
		Task Completion Rate	User task completion rate	≥ 95%	Scenario-based Usability Testing

Objective	Research Question	KPI	Metric	Target	Method of Data Collection
Ensure accessibility	Is the platform accessible to users with disabilities?	Accessibility Compliance	WCAG compliance level	WCAG 2.1 Level AA compliance	Accessibility Audits
Ensure inclusiveness	Can the platform adapt to cognitive abilities, digital literacy and cultural differences?	Inclusiveness Compliance	User customization to personal needs and preferences	≥ 85% rate customization options as "easy to access and use"	Success rate and satisfaction questionnaire
Maintain data security and privacy	Is user data handled securely and in compliance with GDPR?	Security Compliance	GDPR compliance	100% compliance	Security Audits, Data Logs
		Data Processing Accuracy	Accuracy of data handling	≥ 99% accuracy	Data Processing Logs
Ensure ethical AI behaviour and fairness	Is the AI fair, ethical, and transparent in its decisions?	Fairness Compliance	Predictive Parity and Bias Levels	≤ 5% observed bias	AI Fairness Testing Tools, XAI Analysis
		Explainability	User understanding of AI	≥ 75%	Surveys, Transparency Reporting
Validate scalability and resilience	Can the platform handle increased traffic without performance issues?	Scalability Performance	Peak user handling capacity	Full performance under peak load	Load Testing, Stress Simulation
Foster user trust in AI and platform	Do users trust the platform and AI features?	Trust Score	User trust rating	≥ 80% trust score	Trust Surveys, Focus Groups
Ensure reliable data handling	Are data inputs and outputs processed accurately and securely?	Data Processing Reliability	Data accuracy and security logs	≥ 99% accuracy	Data Handling Logs

### 2.3 Testing scenarios

The testing scenarios for the ITHACA platform are developed around core functionalities and anticipated use cases (UCs) within each pilot site. Each scenario is designed to evaluate how well the platform supports civic engagement activities, user interactions, and AI-driven decision-making. This approach ensures that the platform meets the functional requirements of each use case and

that users can interact seamlessly with its key features in real-world contexts. The functionalities have not yet been implemented. The document includes suggestions for these functionalities, along with corresponding testing scenarios provided in the Annex IV. A combination of testing scenarios, selected from various Use Cases, will be implemented. However, it is unlikely that all scenarios will be included.

### 2.3.1 Testing scenarios in relation to potential functionalities

#### *UC1: Citizen Feedback and Decision-Making Participation*

- **Scenario 1: Citizen Feedback Submission:** Users submit feedback on local governance issues through a structured interface. Testing will ensure that users can navigate the feedback form, input their responses, and submit them successfully. This scenario will assess the platform's response time, data processing accuracy, and ease of navigation, ensuring that feedback is captured reliably and efficiently.
- **Scenario 2: Engagement in Decision-Making Processes:** Users participate in decision-making by voting or responding to public opinion polls on civic matters. This scenario will evaluate the user experience for selecting options, submitting votes, and viewing aggregated results. It will also validate data integrity, ensuring votes and responses are accurately recorded and represented in the final output.

#### *UC2: AI-Driven Civic Recommendations*

- **Scenario 3: Personalized Content and Recommendation Delivery:** The platform provides users with tailored content recommendations based on civic interests and past interactions. Testing will involve assessing the accuracy and relevance of recommendations provided to users, as well as AI performance in responding to user actions. The scenario will verify the effectiveness of the recommendation algorithm and ensure the recommendations align with user profiles.
- **Scenario 4: Explanation of AI-Driven Recommendations:** Users receive explanations about why specific recommendations were made. This scenario will evaluate the clarity and comprehensibility of AI explanations, ensuring that users can easily understand the rationale behind the platform's suggestions. Testing will gather feedback on user trust and satisfaction with the explanations provided.

#### *UC3: Community Event and Activity Coordination*

- **Scenario 5: Event Discovery and Registration:** Users browse, discover, and register for local events and activities promoted on the platform by the municipality. Testing will ensure the event discovery interface is intuitive and that the registration process is smooth. This scenario will also assess system notifications for event updates or reminders, ensuring that users receive timely communication.
- **Scenario 6: Event Feedback Collection:** After attending an event, users are prompted to provide feedback on their experience. This scenario will validate the feedback collection process and ensure that post-event surveys are accessible and easy to complete. The platform's ability to analyse feedback for insights and display them to relevant stakeholders will also be tested.

### *UC4: Civic Polls and Surveys*

- **Scenario 7: Participation in Civic Polls:** Users participate in civic polls to express their opinions on local issues. This scenario will test the poll creation, distribution, and response submission features. It will also evaluate data accuracy and reporting, ensuring user responses are collected and analysed correctly.
- **Scenario 8: Results Viewing and Analysis:** Users access results from civic polls and see aggregate insights. This scenario will test the functionality that enables users to view summary statistics and results. Testing will ensure that data is visually clear, interpretable, and accessible, especially for users with lower digital literacy.

### *UC5: Collaborative Civic Projects*

- **Scenario 9: Community Project Proposal Submission:** Users submit project ideas for local community initiatives. Testing will focus on ensuring that users can fill out project proposal forms, attach relevant documents, and submit them without issues. The scenario will assess the platform's ability to handle attachments and validate data integrity.
- **Scenario 10: Project Voting and Prioritization:** Community members vote on submitted project ideas to prioritise initiatives. This scenario will evaluate the voting interface, ensuring users can easily select and submit their preferences. It will also test the platform's capacity to aggregate votes and display priority rankings effectively.

### *UC6: User Onboarding and Support*

- **Scenario 11: New User Onboarding and Registration:** This scenario focuses on the onboarding experience for new users. It will test account creation, profile setup, and tutorial access to ensure new users can quickly get oriented and understand how to use the platform's features. Feedback will be gathered to improve the onboarding process for users with varying levels of digital literacy.
- **Scenario 12: Access to User Support and Help Resources:** Users can access support resources, such as FAQs, tutorials, and live help options. This scenario will evaluate the effectiveness of the help and support interface, ensuring users can quickly find the assistance they need. It will also test response times for live support options, if available.

## **2.4 Pilot-Specific Adjustments**

Each pilot site will have testing adjustments based on local needs and characteristics. In Brasov, Romania, testing scenarios could emphasise language support to accommodate the linguistic diversity of the region; include, for example, Hungarian language as well. In Martin, Slovakia, testing could include, for example, ease of use for users with potentially lower digital literacy along with Roma and older users. These pilot-specific adaptations ensure that the testing scenarios are tailored to the unique demographics and infrastructure of each site.

By focusing on core functionalities and specific user journeys, these scenarios provide a comprehensive approach to validating the ITHACA platform's effectiveness in facilitating civic engagement. Each scenario supports iterative improvements to the platform, ensuring it aligns with the practical needs and preferences of diverse user communities.

## 2.4.1 Testing Scenarios for ITHACA Use Cases in relation to inclusiveness and accessibility

### *UC1: Enhancing Digital Community Participation for an Elderly Citizen*

- **Scenario 1: Navigation and Interface Accessibility**  
An elderly user with limited IT skills attempts to navigate the platform to access community events and election details. This scenario will assess the platform's accessibility features (e.g., larger fonts, simplified navigation) and ease of use, ensuring that the interface is intuitive and supportive of users with low digital literacy.
- **Scenario 2: Voting Participation Process**  
The elderly user participates in a digital community election. This scenario will test the clarity and ease of the voting process, including how the platform provides prompts or instructions. Success criteria will be based on user confidence in completing the voting task independently.

### *UC2: Enhancing Civic Participation for a Visually and Mobility Impaired Citizen*

- **Scenario 3: Screen Reader Compatibility**  
A visually impaired user interacts with the platform using a screen reader to browse civic content and participate in decision-making. This test ensures that all text, buttons, and navigational elements are compatible with screen readers, with smooth access to information.
- **Scenario 4: Speech Recognition and Text-to-Speech Functionality**  
This scenario evaluates the effectiveness of speech recognition and text-to-speech features for users with mobility impairments. The user provides voice input to participate in a poll or survey, and the platform reads back options. Testing will focus on response accuracy and ease of interaction.

### *UC3: Enabling Blind Citizens' Participation in Democratic Processes*

- **Scenario 5: Accessible Voting System with Audio Prompts**  
A blind user participates in a voting or survey process with audio guidance. This scenario tests the accuracy and timing of audio prompts, ensuring they align with user inputs and provide clear instructions throughout the voting process.
- **Scenario 6: Feedback Submission Using Assistive Tools**  
The user submits feedback on public transport issues using text-to-speech or speech-to-text functionalities. This scenario evaluates the platform's support for assistive tools, focusing on error-free submission and data integrity.

### *UC4: Enhancing Accessibility for Users with Language Barriers*

- **Scenario 7: Multi-Language Interface and Content Accessibility**  
A non-native speaker accesses the platform in their preferred language to review community information and participate in discussions. This scenario will assess the effectiveness of language localization, including accurate translation and contextual appropriateness of terms.
- **Scenario 8: Real-Time Translation for Civic Poll Participation**

The user participates in a civic poll available in multiple languages. This scenario tests real-time translation accuracy and whether responses are accurately recorded in the platform's primary language. Testing will also evaluate if language selection is easily accessible and user-friendly.

### *UC5: Enhancing Civic Participation of the Romani Community*

- **Scenario 9: Inclusive Community Engagement Outreach**  
Romani community members access community resources and event announcements. This scenario will assess the platform's accessibility ensuring the interface encourages engagement.
- **Scenario 10: Feedback Collection from Romani Participant**  
Romani users submit feedback on community development projects. Testing will evaluate the clarity of the feedback form, ease of use, and cultural inclusivity, ensuring that all users feel supported in expressing their views.

### *UC6: Implementing Citizen Engagement for Urban Development*

- **Scenario 11: Submitting Opinions on Urban Projects**  
Users view urban development proposals and submit opinions. This scenario tests the platform's ability to display detailed project information in an understandable format and the ease of submitting feedback, verifying that users can voice their opinions effectively.
- **Scenario 12: Accessing Project Updates and Timeline**  
Users could follow ongoing projects, receiving notifications on milestones and updates. This scenario evaluates the platform's notification system, ensuring timely updates, and verifies that project timelines and progress reports are easily accessible. Evaluation will focus only on the functionality and not the content, if this scenario will be chosen to be implemented.

### *UC7: Facilitating Civic Engagement through Direct Communication*

- **Scenario 13: Direct Messaging to City Officials**  
A user communicates directly with a city official (e.g., the mayor) through the platform's messaging feature. This scenario tests message delivery, response tracking, and whether users receive acknowledgment of their messages.
- **Scenario 14: Real-Time Q&A with City Officials**  
The user participates in a live Q&A session with city officials. This scenario evaluates the platform's ability to host some form of communication, the effectiveness of user-to-official engagement, and the clarity of the Q&A interface.

## 2.5 Additional Pilot-Specific Adaptations

In **Brasov, Romania**, where there is a diverse linguistic background, UC4 and UC5 scenarios will have increased emphasis on language support (e.g., include Hungarian). Meanwhile, in **Martin, Slovakia**, UC1 and UC2 will focus more on onboarding and user guidance features, addressing, for example, potential digital literacy barriers among certain community members.

These scenarios ensure that each use case is thoroughly tested to support an inclusive, accessible, and engaging experience for all users, aligning with ITHACA's goals of fostering democratic participation through adaptable and user-centred platform functionalities.

The final testing scenarios cannot be fully defined at this stage, as the platform is not yet available. To address this, we have developed a preliminary set of testing scenarios based on the identified functionalities (Annex IV). These scenarios will be further refined and tailored to align with the Use Case scenarios outlined in this section. The finalised testing scenarios, reflecting the implemented platform and its specific applications, will be documented in deliverables D4.2 and D4.3.

## 2.6 Validation process for Pilot Sites

The validation process for the two pilot sites, Brasov and Martin, is tailored to ensure the ITHACA platform is tested under realistic, site-specific conditions, incorporating regional, demographic, and technical factors.

The validation process follows a multi-stage approach across three testing phases: expert pre-tests, first-iteration prototype testing, and second-iteration real-life testing. Each phase integrates customised adjustments for each pilot site to accommodate their unique user profiles and technological environments. By structuring these phases to reflect each pilot site's specific demographic, geographic, and technical needs, the validation process is set to yield comprehensive insights into the platform's functionality, inclusivity, and adaptability, informing refinements to support broad civic engagement across varied community settings.

### 2.6.1 Pre-Testing with Experts and Wireframes (Phase 0)

In the pre-test phase, wireframes of the platform are evaluated by a selected group of experts from the domains of digital accessibility, civic engagement, and AI ethics. The goal is to gather feedback on the initial user interface (UI) layout, navigation ease, and perceived accessibility. Experts from the Consortium and both Brasov and Martin will simulate interactions with typical platform functionalities (e.g., content posting, voting, and feedback submission) to validate initial design assumptions. Based on this input, platform adjustments will address visual accessibility needs, clarity of information architecture, and inclusivity in functionality descriptions.

### 2.6.2 First Iteration: Prototype Usability and Performance Testing (Phase 1)

- **Performance Testing:** The prototype will also undergo performance assessments, such as response time and error rates, using analytics tools to simulate different user volumes and network conditions. This phase will measure the platform's technical resilience in both pilot sites, addressing Brasov's larger population and Martin's mixed urban-rural digital infrastructure.
- **Usability Testing:** The initial prototype will undergo controlled usability testing with selected participants from Brasov and Martin to assess navigational ease, clarity of the interface, and initial engagement. These tests will involve structured tasks representing typical platform interactions (e.g., voting on community projects or submitting feedback), enabling testers to provide feedback on the intuitiveness and ease of use. Accessibility will be a primary focus, with targeted evaluations of STT (Speech-to-Text) and TTS (Text-to-Speech) functionalities to ensure accessibility for visually impaired users.

### 2.6.3 Second Iteration: Real-Life Testing in Pilot Communities (Phase 2)

- **Scale and Diversity in User Testing:** In this final stage, the ITHACA platform is deployed within both pilot communities under real-life conditions. In Brasov, testing will focus on high-engagement scenarios with a larger, linguistically diverse user base, emphasising the platform’s multilingual capabilities and support for dense network traffic. In Martin, testing will target the user who have already participated in the WP2 activities. This phase aims to test the platform’s scalability, user satisfaction, and overall impact on civic engagement.
- **Community-Driven Interaction and Feedback Collection:** Both pilot sites will implement customised community engagement strategies to gather diverse user feedback. Brasov’s testing will incorporate multilingual feedback forms and surveys to capture user satisfaction across various language groups, while Martin’s testing with users who participated in the WP2 activities.

## 3. Implementation of the framework

### 3.1 Evaluation stages

The process of intervention for the ITHACA platform’s deployment in the pilot sites of Brasov, Romania, and Martin, Slovakia, follows a systematic framework designed to ensure comprehensive preparation, structured deployment, and detailed evaluation. This framework tailors each stage of the implementation to the unique social, technical, and infrastructural characteristics of each pilot site, supporting an effective, data-driven approach to platform validation.

The initial **Preparation Stage** focuses on laying the groundwork for platform deployment, starting with identifying and engaging relevant stakeholders in each community and the participants who participated in the WP2 activities. In both sites, key stakeholders include municipal authorities, community leaders, non-governmental organisations (NGOs) will support in participant recruitment, technical support, and user engagement. Early meetings (before any phase starts) and workshops are held with these stakeholders to align on the platform’s goals, community impact, and roles for each partner, fostering a collaborative approach to platform integration. Local NGOs and community groups are engaged as advocates, helping promote the platform and facilitating user onboarding. Concurrently, the preparation phase involves defining target user groups through demographic analysis, focusing on diverse representation from marginalised groups, older users, and civic organisations. By working with community-based organisations, the project’s recruitment strategy ensures that these underrepresented groups are included in the pilot, promoting a broad base of community input for platform development.

The **Technical Infrastructure Setup** is customised for each site based on an initial assessment of digital infrastructure availability and will be managed remotely by the technical teams in the Consortium responsible for the ITHACA platform (i.e., KT, SIMAVI).

To support user adoption, the **Training and Capacity Building** component involves a series of training sessions tailored to the needs of local staff. These training sessions provide a hands-on introduction to the platform’s interface, accessibility features, and troubleshooting processes. Training is customised for each user demographic, emphasising key civic engagement functions, such as voting, content posting, and feedback submission. Materials are provided in local languages,

and additional support is offered for users with limited digital experience, ensuring that all participants have the foundational knowledge needed to interact confidently with the platform. User manuals will be provided to the pilot sites, and it will be the responsibility of the pilot sites to translate them. To streamline the process only local staff will be directly involved in the tests. Each site will require the involvement of at least three staff members (i.e., recruitment and appointments, conduction, data collection) for each phase.

The **Deployment Stage** begins with a controlled soft launch, allowing for initial testing with a small, representative group of users, including municipal employees and community leaders. Controlled testing enables a careful assessment of the platform's basic functionality, navigation flow, and stability, with a focus on identifying and addressing potential usability issues before the broader deployment. Feedback collected during this phase informs adjustments to the user interface, accessibility settings, and technical performance, particularly concerning load handling and response times in Brasov, where higher user traffic is anticipated.

Following the soft launch, the platform transitions into **Full Deployment and User Onboarding** across both sites. At this stage, the platform is opened to larger participant groups in each community. Onboarding information will be communicated through community boards, municipal channels, and other public access points to facilitate awareness and encourage participation. While engagement is optional for up to 80 participants, the involvement of 20 participants in Phase 2 remains mandatory. Guidance on features such as accessibility tools, including Text-to-Speech (TTS) and Speech-to-Text (STT) functions, ensuring that participants can fully engage with the platform.

To enhance community engagement, a **Comprehensive Outreach Strategy** will be implemented through both digital and physical channels, including local media, social media, and community notifications to raise awareness. Participants from WP2 activities will primarily be used for the mandatory samples, while the optional samples will be gathered through other methods outlined in this strategy.

The **Ongoing Technical Monitoring and Support** process ensures platform stability and optimal performance during deployment. System monitoring tools continuously track uptime, response times, and user interaction patterns, identifying any technical issues or performance bottlenecks. Technical support is provided through a dedicated helpdesk available in all relevant languages, with escalation protocols in place for rapid resolution of critical issues.

In the final **Evaluation & Iteration Stage**, user feedback and technical performance data are systematically analysed to refine the platform and assess its social impact. User feedback is gathered through digital feedback forms, post-use surveys, and focus groups, capturing both quantitative and qualitative insights on usability, accessibility, and perceived impact on civic engagement. Performance data, including user engagement metrics, task completion rates, and platform response times, are analysed to evaluate the platform's technical efficacy. This stage includes interviews with stakeholders to assess changes in civic engagement and overall community response to the platform, identifying areas for improvement. The findings from these analyses guide iterative enhancements, addressing any identified barriers to usability or accessibility, ensuring the platform meets the diverse needs of each site.

Based on feedback and data insights, **Iterative Improvement and Feature Adjustments** are made to enhance platform functionality and user experience. Modifications to the user interface, navigation flow, and accessibility tools are prioritised, and A/B testing may be conducted to validate these improvements. Staff receive updated training as necessary, ensuring they are equipped to support users in engaging with any new or improved features.

As a final validation step, **Scalability Testing** is conducted in both sites to ensure the platform can handle increased usage as broader community engagement grows. This comprehensive validation process ensures that the ITHACA platform not only performs well under real-world conditions but also successfully meets its objectives of enhancing civic participation and accessibility in both Brasov and Martin.

### 3.2 Time Plan for Pilot Activities

The pilot activities for the ITHACA platform are planned to commence in January 2025, following a structured and detailed timeline to ensure effective preparation, deployment, testing, and evaluation across the pilot sites of Brasov, Romania, and Martin, Slovakia. The process begins with stakeholder engagement and coordination, where the pilot site teams will be mobilised to support user recruitment, technical assistance, and outreach. This stage will also define target user demographics, focusing on marginalised communities, older users, and civic organisations, to ensure inclusivity. Recruitment of WP2 previous activities' participants will be sought, while platform customisation will be initiated to incorporate localised language options, accessibility features, and user interface adjustments specific to each site.

In **February 2025, expert testing** will commence using platform wireframes to evaluate interface usability, navigation ease, and accessibility features. This phase will provide crucial feedback on design and functionality, allowing iterative refinements to address potential issues before the first deployment. Adjustments based on expert input will enhance interface layout, content structure, and initial usability, ensuring alignment with user needs and expectations.

The **first iteration** will take place during **Q1 of 2025**, beginning with performance testing in **March**. This testing will evaluate the platform's technical performance, including load management, responsiveness, and stability under simulated conditions, ensuring it is prepared for anticipated user loads. Following performance testing, user recruitment and onboarding will begin in April. Community awareness campaigns could leverage, if needed, both digital and in-person channels, such as local media and events, to encourage widespread participation. Usability and accessibility testing with users will start in May, focusing on navigation ease, accessibility tools like STT and TTS, and feature intuitiveness. Feedback will be collected through surveys, focus groups, and interaction data, with preliminary analysis performed to identify immediate areas for improvement. To ensure robust insights, the first iteration will include at least 10 users per site, with a minimum of five from vulnerable groups. Deliverable D4.2, originally scheduled for March 2025, will be delayed by two months to accommodate platform development delays and to integrate findings from both expert and initial user testing phases.

From **June to August 2025, iterative improvements** will be implemented based on insights from Phase 1 testing. Adjustments will focus on improving navigation flows, addressing accessibility challenges, and enhancing system responsiveness. Community engagement efforts will continue aiming at maintaining high levels of participation and interaction. Performance testing will be

conducted in July to assess platform scalability and stability ahead of the second iteration, ensuring readiness for broader deployment. This testing will simulate high user load conditions to identify and address any remaining performance issues.

The **second iteration** will begin in **August 2025**, incorporating the final version of the platform and targeting larger-scale real-life testing. This phase will include at least 20 users per pilot site and will focus on evaluating real-time civic engagement, feature adoption, and overall user satisfaction. Data will be collected on participation rates, response times, engagement levels, and accessibility tool utilisation. Final customizations will be made based on performance results and user feedback, emphasising multilingual support, accessibility enhancements, and refined user interaction mechanisms.

The **analysis phase for D4.3** will commence in **October 2025**, with the final deliverable submitted in November 2025. This phase will include comprehensive evaluation of user data, focusing on the platform's technical performance, usability, and impact on civic engagement. Quantitative metrics, such as task completion rates and engagement patterns, will be complemented by qualitative insights gathered through focus groups and surveys. The final report will summarise pilot outcomes, document platform refinements, and provide recommendations for scaling the platform to additional communities. This structured approach, supported by defined dependencies such as stakeholder commitment and performance testing prior to user testing, ensures a scientifically rigorous validation process for the ITHACA platform, addressing both technical and social objectives effectively.

### 3.3 Recruitment and Engagement Strategies

The recruitment and engagement strategies for the ITHACA platform pilot sites are designed to ensure a diverse, inclusive, and active participant base while maintaining consistent involvement from key stakeholders. Recruitment efforts will focus on two primary groups: platform users, including representatives from marginalised and vulnerable communities, and key stakeholders, such as municipal authorities, community leaders, and NGOs.

For user recruitment, strategies will leverage a combination of digital and physical outreach methods, tailored to the unique characteristics of each pilot site. Local media campaigns, social media advertisements, and in-person community events will raise awareness and encourage participation. Collaboration with community organisations and NGOs will be sought, if needed, to reaching underrepresented groups, such as elderly users and those with limited digital literacy. Recruitment will start with the people who participated in the WP2 activities.

Engagement strategies will focus on fostering sustained participation through regular communication and ongoing support. For stakeholders, engagement will be maintained through workshops, progress updates, and collaborative discussions on platform implementation and community impact.

These strategies aim to build trust, ensure broad representation, and maintain active participation throughout the pilot activities, aligning recruitment and engagement efforts with the ITHACA platform's objectives of accessibility, inclusivity, and civic engagement.

### 3.4 User Support and Training Procedures

To ensure users are fully equipped to participate effectively in the ITHACA platform pilot activities, a comprehensive user support and training framework will be implemented. This framework is

designed to address diverse user needs, promote platform accessibility, and enhance overall user experience throughout the pilot.

User manuals will be developed to guide participants in understanding and utilising the platform's functionalities (the structure can be found in Annex VIII). The platform must be finalised as a prerequisite for further activities. User manuals will likely be developed by SIMAVI, with pilot sites responsible for translating them into the local languages. A webinar will be organized to provide training, during which the technical teams will guide the pilot sites on how to use the platform effectively. The user manuals will play a key role in supporting Phase 2 evaluation activities. These resources will include step-by-step instructions on navigating the platform, using key features such as accessibility tools (e.g., STT and TTS), participating in discussions, and submitting feedback. The materials will also cover troubleshooting common issues and maximising engagement through the platform's interactive features. To accommodate the multilingual nature of the pilot sites, these manuals will be translated by the respective pilot site teams into Romanian and Slovak, ensuring inclusivity and ease of understanding.

**Technical Support:** A dedicated technical support service will be established to provide timely technical support to users. This service will be accessible via email, where users can report issues, request assistance, and seek clarifications regarding platform use. The service will be staffed by technical support personnel, including representatives from the pilot site teams, who are familiar with the platform's functionalities and technical requirements. Response times will be optimised to ensure users receive prompt resolutions to their queries, fostering confidence in the platform's reliability. The technical support team will be provided by KT and SIMAVI for remote support. One person from each pilot site will be responsible for communicating any issues with the technical support team in Phase 1 and Phase 2.

This multi-faceted approach to user support and training ensures that participants from diverse backgrounds, including vulnerable and digitally inexperienced groups, can fully engage with the platform. By combining comprehensive training materials, readily available technical assistance, and support, the framework provides a robust foundation for user empowerment and effective participation in the ITHACA platform pilots.

## 4. Evaluation Procedures for the Pilots

This chapter outlines the evaluation framework and procedures designed to assess the performance, usability, and impact of the ITHACA platform during its pilot implementation in Brasov and Martin. The evaluation process aims to systematically measure the platform's technical functionality, user experience, and its ability to enhance civic engagement within the pilot communities. Through a combination of qualitative and quantitative methods, the procedures will capture insights from diverse user groups, including vulnerable populations, and identify areas for iterative improvement. The evaluation will be conducted in three phases—expert walkthroughs, initial usability and performance testing with prototypes, followed by real-life deployment of the final platform version—ensuring a comprehensive understanding of its effectiveness and scalability.

## 4.1 Testing with experts (Phase 0): Cognitive walkthrough of platform wireframes

The initial phase of evaluation, referred to as Phase 0, involves a cognitive walkthrough conducted by a group of 5 to 8 experts specialising in usability, accessibility, human-computer interaction (HCI), and platform functionality. This multidisciplinary team will include usability researchers, accessibility specialists, software engineers, and civic engagement experts to ensure a comprehensive assessment of the platform's design. The primary goal of this walkthrough is to systematically evaluate the usability, functionality, and overall user experience of the platform wireframes during their early design stages, identifying potential issues before broader deployment.

The walkthrough will simulate user tasks based on predefined use cases and scenarios, ensuring the platform meets the diverse needs of its target audience, including users with accessibility requirements. Experts will evaluate each interaction step by step, focusing on key questions such as whether users will understand what to do at each step, whether interface elements are visible and prominent, whether the feedback provided by the system is clear and actionable, and whether users can effectively recover from errors.

A structured checklist of evaluation criteria will be provided to guide the experts, covering navigation flow, interface consistency, accessibility features, and alignment with the platform's intended functionalities. Specific emphasis will be placed on assessing critical features such as Speech-to-Text (STT) and Text-to-Speech (TTS) tools and their integration into the platform interface. Additional areas of focus will include the ease of navigating between sections, performing tasks such as voting or submitting feedback, and the intuitiveness of workflows for completing essential actions like registering or posting content. The error prevention and feedback mechanisms will also be scrutinised to ensure error messages are clear and that users can recover effectively from mistakes.

The process will begin with a preparation phase where experts are briefed on the platform's objectives, target user groups, and functionalities. Testing materials, including wireframes, use case scenarios, and task instructions, will be provided to all participants. During the execution phase, experts will independently perform walkthroughs of assigned tasks, documenting any usability challenges, ambiguities, or inconsistencies they encounter. Observers from the evaluation team will record expert feedback and note observations for further analysis. The process will conclude with a debriefing session, where experts will consolidate findings, share insights, and identify priority issues based on their severity and potential impact on user experience.

The expected outcomes of this cognitive walkthrough include a detailed report of usability issues and actionable recommendations for refinement. Feedback from the experts will inform iterative improvements to the wireframes, ensuring that the platform's design is intuitive, accessible, and aligned with user needs before advancing to the prototype testing phase. By leveraging the insights of a diverse group of experts, this early-stage validation will guide user-centred development and ensure the platform is well-prepared for subsequent pilot activities.

### 4.1.1 Procedure for the Cognitive Walkthrough

The procedure is divided into three main stages: preparation, execution, and debriefing, ensuring a comprehensive evaluation of the platform's design and functionality. A template can be found in Annex V.

## 1. Preparation Phase

### Briefing Session:

- Experts are provided with an overview of the platform's objectives, target user groups, and intended functionalities. This includes details about the platform's civic engagement goals, key features such as STT and TTS tools, and accessibility requirements.
- Use cases and user scenarios that align with the platform's goals are shared to provide context for the evaluation.

### Provision of Materials:

- Experts receive the platform wireframes, task instructions, and a checklist of evaluation criteria. The criteria focus on navigation flow, interface clarity, feature discoverability, accessibility, error handling, and feedback mechanisms.
- Testing scenarios are provided to simulate user interactions, including tasks like registering, navigating sections, voting, and submitting feedback.

### Role Assignment:

- Experts are assigned specific tasks or user personas, ensuring coverage of diverse user perspectives, including those of vulnerable and marginalised users.
- Observers from the evaluation team are assigned to monitor and record expert activities and comments.

## 2. Execution Phase

### Task Simulation:

- Experts independently perform the assigned tasks using the wireframes, simulating user interactions step by step.
- For each task, experts address four core questions:
  - Will the user know what to do at this step?
  - Will the user notice the correct interface elements?
  - Will the user understand the feedback provided?
  - Will the user be able to recover from errors?

### Documenting Findings:

- Experts document their observations, noting any usability challenges, ambiguities, or inconsistencies in the interface design.
- Observers record additional insights, including non-verbal cues (e.g., hesitation) and interactions that suggest potential design issues.

### Identification of Barriers:

- Experts highlight barriers to task completion, such as unclear instructions, poor feature discoverability, or inaccessible elements.

- Accessibility specialists evaluate the integration and functionality of tools like STT and TTS, as well as compliance with accessibility standards.

### 3. Debriefing Phase

#### Group Discussion:

- Experts convene for a collaborative session to discuss findings, share insights, and consolidate observations.
- Identified issues are categorised based on severity (e.g., critical, moderate, minor) and type (e.g., usability, accessibility, technical).

#### Recommendations Development:

- Experts propose actionable recommendations for addressing the identified issues, focusing on iterative improvements to navigation, interface design, and feature functionality.
- Prioritisation is based on the potential impact of the issues on user experience, especially for vulnerable user groups.

#### Evaluation Report:

- The evaluation team compiles a comprehensive report summarising key findings and recommendations. This report will guide the refinement of wireframes and inform subsequent development stages.

The cognitive walkthrough will provide a detailed assessment of the wireframes, identifying usability challenges and accessibility gaps. The expert feedback will ensure the platform design aligns with user needs, facilitating an intuitive and inclusive user experience. By systematically addressing identified issues during this phase, the walkthrough will lay a strong foundation for subsequent testing and iterative development.

## 4.2 Controlled performance and usability testing (Phase 1)

Phase 1 focuses on the controlled testing of the ITHACA platform prototype to evaluate its performance, usability, and accessibility under simulated conditions. This phase will be conducted with a limited group of participants, including municipal employees and a selection of vulnerable users, to ensure the platform meets technical and user-centred requirements. The testing will prioritise key metrics and KPIs such as system responsiveness, task completion rates, and accessibility compliance, as defined in the evaluation framework. Insights gained during this phase will guide iterative improvements, preparing the platform for broader deployment in Phase 2.

### 4.2.1 Testing Procedures for platform performance (back-end)

To ensure the ITHACA platform meets its software quality requirements during Phase 1, a series of rigorous testing procedures will be implemented. These tests will cover unit testing, integration testing, and stress testing to evaluate the platform's performance, reliability, and stability under

controlled conditions. The procedures need to be verified by the technical teams that will perform them as well as the annexed procedures (Annex VI).

### *Software Quality Requirements*

The key software quality metrics for Phase 1 align with the requirements defined in WP1, WP2, and WP3. These metrics ensure the platform is reliable, secure, and performant under real-world conditions:

1. **Availability and Uptime:** The platform should achieve 99.5% uptime, minimising unplanned downtime events to less than 5 minutes per incident.
2. **Performance:** Response times should be under 2 seconds, database query execution within 200 ms, and network latency under 100 ms to ensure smooth user interactions.
3. **Scalability:** The system must support at least 500 concurrent users and maintain data throughput of 1,000 transactions per minute during peak loads.
4. **Error Rates:** System errors must be minimised, with no more than three critical errors and ten minor errors per month.
5. **Resource Utilisation:** CPU usage during peak load should remain below 70%, and memory usage should stay below 75%, ensuring no resource bottlenecks.
6. **Data Integrity and Security:** The platform must achieve 100% data accuracy in test scenarios and zero critical security vulnerabilities during penetration testing.
7. **Backup and Recovery:** Full system backup and recovery should be completed within one hour to ensure business continuity.

### *Component (unit) Testing*

Unit testing will focus on verifying the functionality of individual components of the platform, such as user authentication, data input forms, and accessibility tools like STT and TTS. Each module will be tested in isolation to ensure it operates as intended, with test cases covering expected and edge-case scenarios. This approach ensures that each component is robust and functions independently before integration.

### *Integrated Platform Testing*

Integration testing will evaluate how different platform components interact with each other. For example, the integration of accessibility tools (e.g., STT/TTS) with content management systems and the functionality of database queries with user input fields will be tested. The goal is to identify and resolve any issues related to data flow, interface compatibility, or functionality between modules.

### *Security and Stress Testing*

Security testing will focus on identifying vulnerabilities and ensuring the platform is resilient against cyber threats. Penetration testing tools, such as OWASP ZAP, will be used to evaluate the platform's defence mechanisms and ensure zero critical security vulnerabilities.

Stress testing will simulate peak load conditions to evaluate the platform's capacity to handle high volumes of concurrent users and transactions. Tools like LoadRunner will be used to measure metrics such as response time, system uptime, error rates, and data throughput under heavy traffic.

Load testing will assess the platform's ability to support at least 500 concurrent users and process over 1,000 transactions per minute while maintaining optimal performance.

#### 4.2.2 User Acceptance Testing (UAT)

User Acceptance Testing (UAT) for the ITHACA platform is designed to comprehensively evaluate its usability, functionality, and accessibility while gathering valuable feedback from stakeholders and end-users. This critical phase ensures that the platform aligns with the requirements and expectations of its diverse target audience prior to final deployment. UAT will employ structured methodologies, including usability surveys, focus groups, and real-time usage tracking, to assess user satisfaction and the platform's capacity to support participatory democracy.

The UAT process is guided by clear success criteria. For usability, at least 80% of users should rate the platform as "easy to use." Accessibility features, such as STT, TTS, and font customization, must achieve a minimum satisfaction rate of 90% among users with disabilities. Additionally, at least 90% of key platform features, such as voting mechanisms, AI summarization tools, and feedback submission, should perform as intended. Task completion rates should reach at least 95% for essential activities like voting or providing feedback, and at least 75% of registered users should actively engage with the platform on a monthly basis during the testing period.

Key areas of evaluation include usability, which assesses ease of navigation, interaction with features, and overall user experience using tools like usability questionnaires, online surveys, and focus group discussions. Structured instruments such as the System Usability Scale (SUS) or IsoMetrics usability inventory may be adapted for evaluation. Accessibility testing will focus on the functionality and ease of use of features like STT, TTS, and adjustable fonts, with feedback collected through surveys and interviews with users with disabilities. Feature functionality will assess the accuracy and performance of core platform elements, with technical experts conducting initial checks and end-users providing feedback through embedded surveys. Ease of learning will measure how intuitive the platform is for new users through cognitive walkthroughs, post-use surveys, and interviews. User engagement will track active participation, including voting, submitting feedback, and accessing features, while analytics tools will monitor usage patterns. User satisfaction and emotional response will capture overall satisfaction with platform features and perceived value, assessing the platform's impact on participatory behaviour.

Feedback will be collected using several methods. Online surveys embedded within the platform will gather real-time feedback after task completion or upon logging out, featuring Likert-scale questions and open-ended responses. Focus groups with representative users, including marginalised communities, will provide qualitative insights into usability challenges and feature adoption. Semi-structured interviews with stakeholders and selected users will offer personalised insights into the platform's strengths and areas for improvement. Platform analytics will track real-time data on user activity, feature usage, and task completion rates, with heatmaps and interaction tracking offering visual insights into user behaviour. Error reporting mechanisms, such as feedback forms and a built-in ticketing system, will log, categorise, and resolve issues iteratively.

The UAT procedure begins with participant recruitment, engaging a diverse group that includes community members, elderly users, individuals with disabilities, and civic organisations. Participants will receive training sessions and user manuals to familiarise them with the platform's features and objectives. Accessibility features will be configured based on specific user needs. During testing,

participants will engage in scenario-based usability tasks such as voting, navigating content, and submitting feedback. Metrics such as task completion time, error frequency, and user satisfaction ratings will be tracked. Collected feedback and analytics data will be analysed to identify trends and patterns, with findings summarised in structured reports. Based on this analysis, adjustments will be implemented, and the platform will undergo further testing to refine its usability and functionality.

The UAT process is expected to ensure that the ITHACA platform meets the needs of diverse user groups, supports seamless functionality, encourages active engagement, and delivers high user satisfaction. Insights gathered during UAT will guide final iterations of platform development, ensuring a robust, user-friendly, and inclusive platform that effectively fosters participatory democracy.

The procedure is described in Annex VII. The questionnaires can be found in Annex I.

### 4.3 Real-life performance and user experience testing (Phase 2)

Phase 2 represents a critical phase in validating the ITHACA platform under real-world conditions. This stage aims to ensure that the final version of the platform, including any modules or integrations introduced since Phase 1, meets the required standards for usability, functionality, and robustness. The testing procedures will extend and adapt the methodologies applied in Phase 1 to account for the complexities of real-life deployment. Notably, the scope of testing will depend on the components and modules added or modified since Phase 1, requiring the reassessment of these new elements using both unit and integration testing. This iterative process ensures that all platform functionalities, including new integrations, align with the defined software quality metrics and objectives.

#### 4.3.1 Performance Testing

##### *Testing procedures for software quality*

The quality evaluation in Phase 2 will involve testing procedures designed to assess the overall platform performance under realistic conditions. The focus will be on ensuring that all integrated modules function seamlessly and that the platform demonstrates resilience under high usage loads. These procedures will include a detailed examination of both unit and integration testing as well as stress testing, ensuring the platform can handle peak user activity.

##### **Component Testing**

Component testing in Phase 2 will target individual modules, including those carried forward from Phase 1 and any new additions. Examples include the voting module, AI-guided summarisation tools, and accessibility features like STT and TTS. Each component will be tested in isolation to verify its functionality against defined criteria. For modules introduced or updated after Phase 1, the same rigorous testing applied earlier will be replicated, ensuring consistency and reliability. This re-evaluation will help identify and address any issues arising from changes in module design or functionality before integrating them into the full platform.

##### **Integrated Platform Testing**

Integration testing in Phase 2 will evaluate how newly added or updated components interact with existing ones. This stage will focus on ensuring smooth data flow and functional compatibility

between modules, such as the interoperability of AI summarisation tools with content management systems or the seamless execution of database queries triggered by user input. Specific attention will be given to testing integrations introduced after Phase 1, with the applied methods aligning with those used earlier to ensure uniform quality across all modules. Additionally, integration testing will include comprehensive evaluations of data security and user privacy, ensuring compliance with GDPR and other relevant regulations.

### *Security and Stress Testing*

Security testing in Phase 2 will focus on identifying vulnerabilities and ensuring the platform's resilience to cyber threats during live user interactions. Penetration testing will assess the platform's defence mechanisms, particularly those safeguarding new integrations. Emphasis will be placed on verifying data protection mechanisms, including encryption and secure user authentication processes. Stress testing, on the other hand, will simulate real-world peak loads to evaluate the platform's ability to handle heavy traffic. This will include testing for performance under at least 100 concurrent users, with metrics such as response time, error rates, and data throughput assessed to ensure stability and optimal operation. If additional components or functionalities have been integrated, stress testing will apply Phase 1 methodologies to these new features, confirming their scalability and resilience under real-life conditions.

Overall, the testing procedures in Phase 2 will adapt to the evolving platform structure, ensuring that any new modules or integrations undergo rigorous evaluation. This iterative approach ensures that the platform is not only operationally robust but also meets user expectations and project objectives for participatory democracy and inclusivity. The findings from these evaluations will provide a comprehensive foundation for validating the platform's readiness for large-scale deployment.

### **4.3.2 User experience testing**

The user experience testing (UXT) for the final deployment of the ITHACA platform will assess its performance in real-life settings, focusing on its effectiveness, impact on civic participation, and alignment with user expectations. Unlike the controlled testing of Phase 1, this phase evaluates the platform under natural usage conditions, where participants interact with it organically, reflecting their everyday needs and behaviours. The objective is to capture insights into user engagement, emotional responses, and the platform's role in fostering participatory democracy, while also addressing any barriers to continued usage. The testing process will combine quantitative and qualitative methodologies to ensure a comprehensive evaluation.

A minimum of 20 users per pilot site will participate in the testing, including a diverse representation of community members, older individuals, marginalised groups, and civic organisations. At least five participants from each site will belong to vulnerable groups, such as individuals with disabilities, older users with limited digital literacy, or members of economically disadvantaged communities. Recruitment will be conducted in collaboration with local NGOs and stakeholders to ensure inclusivity and diversity in the participant pool.

Participants will use the platform over a two-week period, during which their interactions will be monitored through analytics embedded within the platform. Online questionnaires, accessible via links integrated into the platform, will be distributed to capture user feedback periodically, focusing on metrics such as satisfaction, ease of use, accessibility, and perceived value in supporting civic engagement. Surveys will include Likert-scale questions and open-ended responses, ensuring a balance between structured data collection and qualitative insights. Additionally, focus groups with

a subset of users will provide an opportunity to explore user experiences in depth, focusing on challenges, successes, and areas for improvement. These focus groups will target at least five participants per site, including representatives from vulnerable groups, ensuring their perspectives are prioritised.

Interviews with key stakeholders, such as municipal authorities and community leaders, will further complement the evaluation, providing insights into the platform's broader impact on civic engagement and decision-making processes. These semi-structured interviews will explore the effectiveness of the platform's features, its integration into existing workflows, and its potential to support participatory democracy at the community level.

The evaluation will assess various metrics, including engagement levels, retention rates, perceived value, and inclusivity. Emotional responses will be measured through sentiment analysis of feedback and survey responses, capturing feelings of trust, empowerment, or frustration. Additionally, the impact of the platform on civic participation will be evaluated by analysing self-reported actions influenced by the platform, such as participation in community initiatives or engagement with local authorities.

By combining online questionnaires, focus groups, and stakeholder interviews, the user experience testing will provide a holistic understanding of how the platform performs in real-life settings. Insights gathered will guide final refinements, ensuring the platform meets its objectives of inclusivity, usability, and support for participatory democracy. This phase will also validate the platform's scalability and adaptability, preparing it for potential deployment in broader contexts beyond the pilot sites.

The procedure is described in Annex VII and the evaluation material in Annex I.

## 5. Risk assessment and Mitigation

### 5.1 Identified Risks for Pilot Implementation

The successful implementation of the pilot phase for the ITHACA platform depends on identifying and addressing potential risks that could disrupt timelines, hinder user engagement, or compromise the platform's technical integrity.

#### 5.1.1 Technical Challenges in Platform Functionality

Issues such as system errors, slow response times, or failed integrations of core features (e.g., STT/TTS and voting mechanisms) may arise during deployment. The likelihood of these challenges is medium, but their potential impact is high, as disruptions in platform functionality could lead to a poor user experience and diminished trust among participants.

#### 5.1.2 Low User Engagement

Recruitment challenges, lack of interest among target user groups, or insufficient motivation to actively participate in pilot activities could result in low engagement levels. This risk is assessed as medium likelihood and high impact, as insufficient participation would compromise the reliability of evaluation results and limit valuable feedback for platform improvement.

### 5.1.3 Delays in User Recruitment

Recruitment efforts may face delays due to insufficient outreach, limited access to marginalised groups, or logistical challenges in organising onboarding sessions. This risk is medium in likelihood and impact, as recruitment delays could disrupt the testing schedule and reduce time available for iterative adjustments.

### 5.1.4 Inadequate Accessibility for Vulnerable Groups

Accessibility features, such as screen readers or font adjustments, may not fully meet the needs of users with disabilities, leading to dissatisfaction and disengagement. This risk is assessed as low likelihood but high impact, as failure to meet accessibility needs could alienate vulnerable groups and undermine inclusivity goals.

### 5.1.5 Data Security and Privacy Issues

The platform may encounter vulnerabilities in data handling, potentially exposing user information to unauthorised access. Although this risk is low in likelihood, it carries a high potential impact, as security breaches would damage user trust and pose significant legal and ethical challenges.

### 5.1.6 Delays in Platform Development and Customization

Unforeseen technical delays in finalising platform features or customising the platform for pilot sites could impact timelines. This risk has medium likelihood and medium impact, as delays could compress the time available for testing and iterative improvements.

### 5.1.7 Logistical Challenges in Focus Groups and Interviews

Scheduling conflicts, lack of appropriate venues, or technical issues during virtual sessions could disrupt the collection of qualitative insights. This risk is medium in likelihood and impact, as such disruptions may limit valuable feedback for platform evaluation.

## 5.2 Mitigation Strategies

### 5.2.1 Technical Challenges in Platform Functionality

To address potential technical issues, thorough pre-deployment testing, including stress and integration testing, will be conducted to identify and resolve problems. A dedicated technical support team will monitor real-time errors during pilot activities and implement contingency plans to roll back problematic updates.

### 5.2.2 Low User Engagement

Targeted outreach strategies will be developed in collaboration with local NGOs and community leaders to enhance recruitment and participation. Incentives, such as certificates of contribution or small rewards, could be provided to encourage engagement, but this is a decision to be made by each pilot site team separately. Continuous communication through regular updates and engagement activities will sustain interest.

### 5.2.3 Delays in User Recruitment

Recruitment efforts will begin well in advance and leverage multiple channels, including social media, local events, and direct outreach. A backup plan will involve engaging additional users to quickly reach target demographics if recruitment falls short.

### 5.2.4 Inadequate Accessibility for Vulnerable Groups

Accessibility experts and representatives from vulnerable groups will be engaged early in the development process to test and refine features. Clear tutorials and helpdesk support for accessibility tools will ensure effective usage and address any gaps identified during testing.

### 5.2.5 Data Security and Privacy Issues

Regular security audits and penetration tests, using tools like OWASP ZAP, will ensure the platform's compliance with GDPR and other regulations. User data will be anonymized and encrypted, and privacy policies will be clearly communicated to all participants to build trust.

### 5.2.6 Delays in Platform Development and Customisation

Regular coordination meetings with the technical team will track progress and resolve bottlenecks. Incremental releases will be prioritised to avoid affecting the entire timeline, ensuring critical features are ready on time.

### 5.2.7 Logistical Challenges in Focus Groups and Interviews

Flexible scheduling tools will coordinate participant availability, and backup options, such as virtual platforms, will be tested and prepared in advance. Clear instructions will be provided to participants and moderators to minimise disruptions and ensure successful data collection.

## 6. Ethics and Data Privacy

### 6.1 Ethics requirements and consent

The pilot implementation of the ITHACA platform will adhere to strict ethical standards to ensure the protection and respect of all participants. Ethical considerations include obtaining and managing informed consent, ensuring transparency, and safeguarding participants' autonomy throughout the study. Informed consent will be obtained from participants through a structured process. Before engaging in the pilot activities, participants will be provided with detailed information about the study objectives, their expected roles, the nature of the data to be collected, and how their feedback will contribute to the platform's development. This information will be delivered via a Participant Information Sheet (PIS) in a language accessible to the participants, ensuring inclusivity.

Consent will be collected using signed or digitally recorded forms, depending on the participant's preference and the nature of the interaction (e.g., in-person or virtual). The consent form will explicitly outline participants' rights, including the right to withdraw at any stage without any repercussions and the assurance that their data will be used exclusively for the purposes outlined in the PIS. Special

care will be taken to ensure participants from vulnerable groups, such as individuals with disabilities or low digital literacy, fully understand the consent process. This will involve offering additional explanations or assistance where necessary. The consent process will comply with international ethical guidelines and relevant national laws governing research and data collection.

## 6.2 Data privacy and GDPR compliance

Ensuring data privacy and protection is a fundamental aspect of the pilot activities, with strict adherence to GDPR and other relevant privacy regulations. All data collection, storage, processing, and analysis protocols have been designed to safeguard participant information. Personal data collected during the study, such as demographic details and user interactions with the platform, will be anonymized or pseudonymized wherever possible to protect individual identities.

Data will be securely stored on encrypted servers, with access restricted to authorised personnel involved in the project. User data will be handled with transparency, and participants will be informed about the specific types of data being collected, the purpose of its collection, and how it will be used. No data will be shared with third parties without explicit consent from the participants.

The platform will implement robust data security measures, including regular audits, penetration testing, and real-time monitoring, to prevent unauthorised access or breaches. Backup protocols will be established to ensure data integrity, and contingency plans will be in place to address any potential security incidents.

In compliance with GDPR, participants will have the right to access, rectify, or delete their personal data at any point during the study. The designated Data Protection Officer (DPO) or relevant officer will oversee compliance and address any concerns raised by participants. By embedding these measures into the pilot implementation, the project team ensures a high standard of ethical and data privacy practices, fostering trust and confidence among participants and stakeholders.

# 7. Conclusion and Next Steps

## 7.1 Summary of the Validation Framework

The validation framework for the ITHACA platform has been meticulously designed to ensure a comprehensive and methodical approach to assessing the platform's technical performance, usability, accessibility, and user experience. The framework is structured across multiple phases, starting with expert evaluations of wireframes (Phase 0), followed by controlled performance and usability testing of the prototype (Phase 1), and culminating in real-life performance and user experience testing with the final version of the platform (Phase 2). Each phase incorporates rigorous testing procedures, such as software quality assessments, stress and security testing, and user-centred evaluations through surveys, focus groups, and interviews. A strong emphasis has been placed on inclusivity, involving diverse user groups, including individuals with disabilities, elderly users, and those with limited digital literacy. The framework not only ensures robust technical validation but also examines the platform's ability to foster civic engagement and participatory democracy, aligning with the project's overarching objectives.

## 7.2 Recommendations for Pilot Implementation

To enhance the success of the pilot implementation, several practical recommendations have been identified:

- **Strengthen Recruitment Strategies** by early engagement with local NGOs, municipal authorities, and community leaders will be critical to ensuring diverse and representative user participation. Special attention should be given to reaching vulnerable groups to align with inclusivity goals.
- **Ensure Technical Readiness** by conducting thorough pre-deployment testing to address potential technical challenges. Establish a dedicated technical support team (from KT and SIMAVI) to resolve issues promptly during the pilots.
- **Enhance Training and Support** by providing clear and accessible user manuals. Offer continuous support through the technical support system to assist participants throughout the testing period.
- **Prioritise Accessibility** by collaborating with accessibility experts to refine tools such as STT, TTS, and customization options. Regularly test these features with target user groups to ensure they meet specific needs.
- **Refine Data Collection Protocols** by using anonymised and aggregated data wherever possible to maintain user privacy. Clearly communicate the purpose and scope of data collection to participants to foster transparency and trust.
- **Monitor and Adapt** by establishing mechanisms for ongoing feedback collection during the pilots to identify and address issues in real-time. Adjust timelines and resources as necessary to accommodate unforeseen challenges.

## 7.3 Future Work and Follow-up Activities

Following the completion of the pilots, the focus will shift to analysing the collected data and refining the platform based on user feedback and performance insights. Detailed analysis reports will be prepared, summarising the outcomes of the validation phases and providing actionable recommendations for platform improvements. These findings will be documented in upcoming deliverables, such as D4.2 and D4.3, which will cover the results from Phase 1 and Phase 2 evaluations, respectively.

Future work will include adjustments to platform features and functionalities, particularly those related to accessibility, usability, and scalability. Additional validation may be necessary for newly implemented or updated components. A roadmap for broader deployment of the platform across other communities will be developed, leveraging the insights gained during the pilots.

Finally, dissemination activities, including stakeholder workshops, public summaries, and academic publications, will share the project's outcomes and foster dialogue on the role of digital platforms in supporting participatory democracy. These efforts will ensure the ITHACA platform's scalability and long-term sustainability as a tool for empowering communities and enhancing civic engagement.

## References

CERTH, UniGraz, & UPAT. (2023). Study on good practices of citizen engagement and democracy in AI applications (D1.1). ITHACA Project, European Union Horizon Europe Framework Programme (Grant Agreement No. 101094364).

UniGraz, UPAT, SnP, CERTH, & others. (2023). Trustworthy AI compliance practices, assessment, and conceptualization (D1.3). ITHACA Project, European Union Horizon Europe Framework Programme (Grant Agreement No. 101094364).

RtF, UniGraz, BMA, PEDAL, MARTIN, SnP, MNL, & CERTH. (2024). Report on Citizen Jury Process and User requirements (D2.1). ITHACA Project, European Union Horizon Europe Framework Programme (Grant Agreement No. 101094364).

# Annexes

## Annex I: Pilot Site Details

### Pilot Site 1: Brasov, Romania

#### 1. Pilot Site Overview

- **Location:** Brasov, Romania. The city has a diverse urban environment with modern infrastructure in its central areas and historical sites, making it a suitable testbed for public engagement technology.
- **Site Characteristics:**
  - **Urban Setting:** Brasov's infrastructure includes robust Wi-Fi availability in public buildings, reliable mobile networks, and a variety of public spaces suitable for both digital and in-person community engagement activities.
  - **Community Context:** Brasov's population spans various socio-economic backgrounds, with a mix of active civic organisations, schools, government offices, and public service centres that are familiar with digital platforms. However, certain groups, such as older adults and economically disadvantaged populations, have varying levels of digital literacy.
  - **Digital Engagement Readiness:** Brasov has a relatively high adoption rate for digital services, which will facilitate pilot acceptance, though disparities exist in digital access, which may affect engagement among marginalised groups.
- **Primary Goals:**
  - **Inclusion and Access:** Provide tools that enhance digital accessibility for marginalised communities, including older adults, low-income households, and minority populations.
  - **Trust and Transparency in AI:** Strengthen public trust in AI technology by prioritising transparent processes, especially related to data handling and privacy.
  - **Community-Driven Impact:** Assess the platform's effectiveness in increasing local governance engagement, with metrics on participation, diversity in user demographics, and feedback from community stakeholders on civic engagement quality.

#### 2. Demographic Information

- **Target User Groups:** Diverse user base including:
  - **Older Residents:** Focused efforts on accessibility tools like TTS (Text-to-Speech) and simplified navigation features to support digital inclusion.
  - **Minority Communities:** Resources and language support to engage Hungarian-speaking populations, including cultural sensitivity in platform content.
  - **Youth Organizations and NGOs:** Active users expected to engage more frequently with content creation and voting features, providing detailed feedback on usability and community impact.
  - **Local Government Officials:** Municipal employees from the pilot site team will participate, to conduct the tests.
- **Population Size:** Estimated optimal total of 100 participants (considering all phases).
  - **Phase 0: Q1 early 2025:** Expert testing through cognitive walkthroughs and low fidelity prototypes or wireframes.
  - **Phase 1 (Q1 later 2025):** Approximately 10 participants, mainly pilot staff, municipal employees, and NGO representatives for controlled small-scale usability testing.
  - **Phase 2 (Q3 later 2025):** Expanded testing with an additional 20-30 participants, incorporating the general public to test platform scalability and broader community engagement.
- **Key Demographic Factors:**

- **Language Diversity:** Primary language is Romanian, with secondary support for Hungarian communities.
- **Digital Literacy Levels:** Varied, with notable gaps among older adults and low-income populations.
- **Accessibility Needs:** Emphasis on STT (Speech-to-Text) and TTS integration, large font options, and simplified navigation modes for older adults and those with disabilities.

### 3. Technical Infrastructure

- **Platform Setup:**
  - **Private Device Compatibility:** Ensure compatibility across mobile and desktop devices, including older hardware with limited memory and lower processing capabilities.
- **Technology Accessibility:**
  - Device support could be provided for those who lack personal access, with provision of 2-3 laptops available at pilot team offices.
  - **Training and Assistance:** Comprehensive user guides will be provided for the Phase II evaluation. A dedicated support email from the technical platform team will address user issues during initial deployment and ongoing use.
- **Integration Requirements:**
  - **Compliance Protocols:** Ensure integration processes on ITHACA platform meet national data protection regulations, particularly GDPR, focusing on secure handling of user inputs and anonymized storage for analysis.

### 4. Pilot Implementation Plan

- **Expected Timeline:**
  - **Q1 2025 (Phase 0 and 1):** Controlled testing for platform performance and usability, with small-scale testing among municipal employees, selected NGO members, and representatives from each target demographic group.
  - **Q3 2025 (Phase 2):** Larger community-focused pilot, with system refinements based on Phase 1 feedback. Includes real-life testing with at least 20-30 participants.
- **Recruitment & Engagement Strategies:**
  - **Digital Campaigns:** Collaborate with city council and local NGOs to reach target user groups through social media, local websites, and newsletters.
  - **On-Site Recruitment Events (as a suggestion only for the optional recruitment of the additional 80 participants in Phase 2):** Host community events in public locations to introduce the platform, including hands-on demos and registration assistance.
  - **NGO Partnerships:** Partner with NGOs that work with underserved groups to facilitate recruitment, focusing on inclusivity and ensuring all community voices are represented (as conducted in WP1 and WP2).
- **Training and Support:**
  - **Digital Literacy Workshop:** Offer introductory workshop where users can learn the basics of platform navigation, account setup, and specific features like voting or feedback submission.
  - **Technical support team:** Establish a dedicated email for local staff to receive assistance with technical issues.
  - **Localised User Guides:** Provide printed and digital guides in Romanian and Hungarian, with visuals to support users with limited literacy.

### 5. Evaluation and Data Collection

- **Evaluation Methods:**
  - **Phase 0:** Expert (cognitive) walkthroughs
  - **Phase 1:** Usability tests (task completion rates, time taken per task, error tracking) and performance evaluations (response times, error rates).

- **Phase 2:** Detailed surveys, interviews, and focus groups to capture qualitative data on accessibility, inclusivity, and overall satisfaction, alongside quantitative data on usage patterns and feature engagement.
- **Data Collection Tools:**
  - **Digital Logs:** Collect user interaction logs to track engagement metrics.
  - **Feedback Forms and Surveys:** Surveys on ease of use, accessibility, and trust in AI functions.
  - **Focus Groups:** Semi-structured discussions with a diverse sample of users to gather feedback on specific pain points and suggestions for improvement.
- **Key Metrics:**
  - **Engagement Rates:** Frequency of interactions per user, content creation, and responses.
  - **User Satisfaction Ratings:** Overall satisfaction with platform features and ease of use.
  - **Accessibility and Inclusivity:** Use of accessibility features (STT, TTS), success rates among older users, and engagement of marginalised groups.

## 6. Risks and Challenges

- **Potential Risks:**
  - **Digital Literacy Barriers:** Risk of low adoption by the municipality or among older adults and economically disadvantaged users.
  - **Privacy Concerns:** Potential resistance from users concerned about data security, especially regarding AI data processing.
- **Mitigation Strategies:**
  - **Clear Data Transparency Policies:** Explicit user consent agreements and clear explanations of data usage will be provided to alleviate privacy concerns.
  - **Emphasizing Benefits for Municipalities:** Highlight tangible advantages, such as improved decision-making through data-driven insights, increased public trust, and enhanced reputation for fostering inclusivity. Demonstrate how participatory governance can streamline resource allocation, identify community priorities efficiently, and reduce opposition to municipal policies by involving residents in the decision-making process.

## 7. Additional Information

- **Local Partnerships:** Collaboration with the Brasov city council and/or other city councils from metropolitan area, local community organisations, and NGOs for participant outreach, training, and feedback collection.
- **Site-Specific Considerations:** Adherence to local and EU data protection laws, as well as platform localisation to support Hungarian-speaking users, are vital for platform adoption and trust.

## Pilot Site 2: Martin, Slovakia

### 1. Pilot Site Overview

- **Location:** Martin, Žilinský kraj, Slovakia, a smaller urban area with a tightly knit community and a mix of public institutions, community organisations and civic groups.
- **Site Characteristics:**
  - **Urban-Rural Blend:** Martin offers a unique testing environment with both urban and rural characteristics, which will provide insights into the platform's adaptability to diverse community settings.
  - **Community Context:** Martin has a well-established social infrastructure with active local governance, schools, community centres, and cultural institutions, making it an ideal setting for testing civic engagement technologies.
- **Primary Goals:**
  - **Improve Civic Engagement:** Encourage Martin's residents to engage in community discussions and decision-making through a user-friendly digital platform.

- **Enhance Digital Literacy:** Test the platform's effectiveness as a tool to increase digital literacy among older and low-tech users.
- **Test Platform Scalability in Rural Contexts:** Assess the platform's performance and relevance in both urban and rural settings.

## 2. Demographic Information

- **Indicative User Groups:**
  - **Older Residents:** Specific focus on providing accessible features such as large-font text, audio navigation through TTS, and simplified navigation.
  - **Youth and Community Groups:** Engage younger populations and community organisations who are generally more tech-savvy and may act as community influencers.
  - **Government and Local Authorities:** Participation from local officials and community leaders to test the platform's utility in disseminating information and gathering public input.
  - **Low-Digital Literacy Individuals:** Tailored support for individuals with limited experience in digital platforms, ensuring inclusivity.
  - As participants, **other groups** to consider are:
    - Younger / youth (i.e. 18-30)
    - Refugees & Migrants
    - Romani
    - Woman
- **Population Size:** Approximately and optimally 80-100 (participants for all phases, but mostly for real-life use (Phase 2)).
  - **Phase 0: Q1 early 2025:** Expert testing through cognitive walkthroughs and low fidelity prototypes or wireframes with Consortium partners.
  - **Phase 1 (Q1 later 2025):** Small-scale testing with an initial group of 10 municipal employees, users and community leaders.
  - **Phase 2 (Q3 later 2025):** Expanded testing to 20 general community members, targeting a diverse demographic to evaluate full platform scalability. The rest of 80 to reach 100 participants are optional.
- **Key Demographic Factors:**
  - **Language:** Slovak is the primary language, and platform resources will be provided in Slovak to ensure local relevance.
  - **Digital Literacy Variability:** A significant portion of participants will be lower digital users.
  - **Accessibility Needs:** Key accessibility features like STT and TTS, adjustable text size, and high-contrast modes will be prioritised for older adults and users with visual impairments.

## 3. Technical Infrastructure

- **Platform Setup:**
  - **Device Compatibility:** Ensure compatibility across mobile devices and computers, including those with lower specifications, to avoid accessibility issues among participants with limited device options.
- **Technology Accessibility:**
  - **Device Lending Programme:** Could establish a device lending programme in collaboration with local community centres to provide tablets or laptops for users who lack personal access (only for Phase 2).
  - **Technical Support Staff:** Appoint technical teams' representatives (from KT and SIMAVI) to provide technical support via email and assist local staff with issues, problems, etc.
- **Integration Requirements:**
  - **Compliance and Data Protection:** Ensure that platform use complies with national data protection standards, especially when handling citizen feedback and demographic data.

#### 4. Pilot Implementation Plan

- **Expected Timeline:**
  - **Q1: early 2025 (Phase 0):** Expert (cognitive) walkthroughs.
  - **Q1 later 2025 (Phase 1):** Initial usability and platform performance testing focused on platform stability and interface usability.
  - **Q3 2025 (Phase 2):** Expanded pilot testing with larger community groups, covering usability, engagement, and data-driven feedback collection for broader civic involvement.
- **Recruitment & Engagement Strategies:**
  - **Digital Campaigns:** Collaborate with city council and local NGOs to reach target user groups through social media, local websites, and newsletters.
  - **Partnership with Local Organisations:** Collaborate with Martin's local NGOs, schools, and cultural institutions to leverage existing networks for recruitment, emphasising inclusivity.
- **Training and Support:**
  - **Digital Literacy Workshop:** Offer introductory workshop where users can learn the basics of platform navigation, account setup, and specific features like voting or feedback submission.
  - **Technical support team:** Establish a dedicated email for local staff to receive assistance with technical issues.
  - **Step-by-Step Guides:** Provide printed and digital guides in Slovak, tailored for low-tech users with visual instructions and clear, simple language.

#### 5. Evaluation and Data Collection

- **Evaluation Methods:**
  - **Phase 0:** Expert (cognitive) walkthroughs.
  - **Phase 1 (Usability and Performance Tests):** Evaluate the platform's performance with a controlled group by tracking key metrics such as response time, error rate, and user feedback on navigation ease.
  - **Phase 2 (Community Engagement Testing):** Full-scale evaluation with diverse participants to assess inclusivity, accessibility, and user satisfaction. Includes feedback collection through surveys, focus groups, and individual interviews.
- **Data Collection Tools:**
  - **Digital Tracking:** Use analytics tools to monitor user engagement, feature usage, and navigation patterns.
  - **Surveys and Feedback Forms:** Distribute usage surveys to gather insights on user experience, accessibility, and impact on civic engagement.
  - **Focus Groups and/or Interviews:** Conduct focus groups with different user groups to capture in-depth feedback on platform usability, accessibility features, and perceived impact.
- **Key Metrics:**
  - **Usability Scores:** Track user ratings on ease of navigation, accessibility, and overall platform satisfaction.
  - **Engagement Levels:** Measure frequency of interactions, time spent on the platform, and diversity in user-generated content.
  - **Accessibility and Inclusivity Metrics:** Gauge the usage of accessibility features (e.g., STT, TTS) and assess platform's inclusivity based on demographics like age and digital literacy.

#### 6. Risks and Challenges

- **Potential Risks:**
  - **Low Device Access:** Limited access to digital devices among low-income users could hinder platform adoption.
  - **Digital Literacy Barriers:** High likelihood of low engagement from users unfamiliar with digital technology, requiring extensive onboarding support.

- **Data Privacy Concerns:** Potential concerns about personal data usage and security, especially among older residents.
- **Mitigation Strategies:**
  - **Device Lending and Support Programmes:** Could provide device access through community centre programs and ensure on-site support for those with limited experience, if such need is relevant and arises.
  - **Privacy Assurance:** Communicate clear and transparent data handling policies, including user consent and anonymisation practices, to build trust.

## 7. Additional Information

- **Local Partnerships:**
  - Work closely with Martin's municipal offices, NGOs, and cultural institutions to facilitate recruitment, engagement, and data collection. These organisations can help ensure that a broad spectrum of the community is represented in the pilot.
- **Site-Specific Considerations:**
  - Adherence to Slovak data protection laws will be crucial, with additional sensitivity to local community preferences around privacy and data security.
  - Special attention will be given to minority and vulnerable populations, ensuring that all user groups feel supported and included in the digital engagement process.

## Pilot Site Description Template

### 1. Pilot Site Overview

- **Location:** (e.g., City, Region, Country)
- **Site Characteristics:** Briefly describe the site's key characteristics (e.g., urban or rural, population size, local infrastructure).
- **Primary Goals:** Outline the main objectives for implementing the pilot in this site.

### 2. Demographic Information

- **Target User Groups:** Describe the specific user groups that will participate (e.g., age range, community stakeholders, accessibility needs).
- **Population Size:** Indicate the approximate number of participants expected.
- **Key Demographic Factors:** Note any relevant demographic factors, such as predominant languages, socio-economic characteristics, or specific accessibility needs.

### 3. Technical Infrastructure

- **Platform Setup:** Describe any existing infrastructure relevant to the pilot (e.g., available connectivity, hardware/software needs).
- **Technology Accessibility:** Indicate if the site has any unique technological considerations, such as connectivity limitations or device compatibility issues.
- **Integration Requirements:** Identify any specific integration needs (e.g., integration with local systems, databases, or third-party applications).

### 4. Pilot Implementation Plan

- **Expected Timeline:** Provide an estimated timeline, including key milestones from deployment to completion.

- **Recruitment & Engagement Strategies:** Describe the planned strategies to recruit and engage users (e.g., outreach programs, local partnerships, community events).
- **Training and Support:** Outline how training and user support will be provided to participants. Mention whether training sessions, guides, or other resources will be offered.

## 5. Evaluation and Data Collection

- **Evaluation Methods:** List the evaluation methods that will be used (e.g., surveys, interviews, focus groups).
- **Data Collection Tools:** Specify the tools or instruments that will be used to collect data (e.g., feedback forms, digital logs).
- **Key Metrics:** Identify any specific metrics that the site will focus on, such as user engagement, accessibility, or system performance.

## 6. Risks and Challenges

- **Potential Risks:** Outline any anticipated challenges or risks related to implementing the pilot in this site (e.g., technical, logistical, or user-related risks).
- **Mitigation Strategies:** Describe strategies to mitigate each identified risk.

## 7. Additional Information

- **Local Partnerships:** Mention any local partnerships or stakeholders involved in supporting the pilot.
- **Site-Specific Considerations:** Note any additional site-specific factors that could impact the pilot, such as local policies, regulatory considerations, or cultural factors.

## Annex II: Evaluation Forms and Questionnaires

The questionnaires will be prepared using the Typeform platform, with links integrated into the ITHACA platform. This approach will be fully implemented for the Phase 2 evaluation, where the online questionnaires will be administered to participants.

### Phase 1: Controlled usability testing

#### Usability Questionnaire

**Purpose:** Collect user feedback on the platform's ease of navigation and interaction, focusing on task completion and overall user experience.

#### Questions

1. How easy was it to navigate the platform? (Likert scale: Very Difficult - Very Easy)
2. Were you able to locate the features you needed quickly? (Yes/No; If no, specify the difficulty)
3. How intuitive did you find the process of completing tasks, such as submitting a vote or posting content? (Likert scale)

4. Were there any sections or features of the platform that felt confusing or unclear? (Open-ended)
5. On a scale from 1 to 5, how would you rate the overall usability of the platform? (1 = Poor, 5 = Excellent)

## Accessibility Questionnaire

**Purpose:** Evaluate the performance of accessibility features such as STT, TTS, and font customization, ensuring inclusivity for users with disabilities.

### Questions

1. Did the platform's accessibility features meet your needs? (Yes/No; If no, explain why)
2. How easy was it to use the Speech-to-Text (STT) feature? (Likert scale)
3. How satisfied were you with the Text-to-Speech (TTS) functionality? (Likert scale)
4. Were you able to adjust the fonts, colours, or sizes to suit your preferences? (Yes/No)
5. Do you have any suggestions for improving the accessibility of the platform? (Open-ended)

## User Satisfaction Survey

**Purpose:** Assess overall user satisfaction with the platform's features, usability, and performance.

### Questions

1. How satisfied were you with the overall experience of using the platform? (Likert scale)
2. Which features of the platform did you find most useful? (Multiple choice; with an "Other" option for specifics)
3. Were there any features that you found unnecessary or difficult to use? (Open-ended)
4. On a scale from 1 to 5, how likely are you to recommend this platform to others? (1 = Very Unlikely, 5 = Very Likely)
5. How well do you think the platform supports civic engagement and participatory democracy? (Likert scale)

## Focus Group Discussion Guide

**Purpose:** Facilitate in-depth qualitative feedback from selected users, focusing on usability, accessibility, and the platform's role in civic engagement.

### Discussion Topics

1. **First Impressions:** What were your initial thoughts about the platform's layout and features?
2. **Ease of Use:** Were there any tasks you found particularly easy or difficult to complete? Why?
3. **Accessibility:** Did the accessibility features help you effectively use the platform? What could be improved?
4. **Civic Engagement:** Do you feel the platform supports your ability to participate in community decision-making? If so, how? If not, why?

5. **Suggestions for Improvement:** What would you change or add to the platform to make it more effective?

## Semi-Structured Interview Questions

**Purpose:** Gather personalised feedback from key stakeholders and select users on the platform's performance, features, and areas for improvement.

### Questions

1. Can you describe your overall experience using the platform?
2. Were there specific features that stood out to you, either positively or negatively?
3. How effective do you think the platform is in supporting civic engagement and democracy?
4. Did you encounter any technical issues or usability challenges? If so, what were they?
5. What improvements would you suggest making the platform more user-friendly and impactful?

## User Feedback Forms

**Purpose:** Enable users to report issues and suggest improvements during platform usage.

### Questions

1. Describe any issues or bugs you encountered while using the platform. (Open-ended)
2. What features or functions would you like to see added or improved? (Open-ended)
3. How would you rate your overall experience with the platform today? (Likert scale)
4. Do you have any additional comments or suggestions? (Open-ended)

## Data Collection Methodology

- **Online Questionnaires:** Embedded surveys will be distributed post-usage, allowing users to provide structured feedback anonymously.
- **Focus Groups:** Conducted with 5+ participants per pilot site to gain qualitative insights.
- **Interviews:** Involving 5+ stakeholders per site to collect detailed individual feedback.
- **Automated Platform Analytics:** Monitoring user behaviour, task completion rates, and engagement metrics.

## Phase 2: Real-life user experience testing

### Before Platform Interaction: Pre-Interaction Survey

The pre-interaction survey should focus on:

- **Baseline knowledge** of civic engagement platforms.
- **Expectations** for the platform.
- **Experience** with similar tools.

## Pre-Interaction Survey Questions

### Demographics and Background

1. Have you ever used a digital platform for civic participation or community engagement? (Yes/No)
  - If Yes, which one(s)?
2. How comfortable are you with using digital platforms? (Scale: 1 - Not Comfortable, 5 - Very Comfortable)
3. What is your primary device for accessing digital tools? (Options: Smartphone, Tablet, Desktop, Other)

### Expectations for the Platform

1. What features would you expect in a platform designed to support community engagement? (Open-Ended)
2. How do you expect this platform to help with your civic roles or responsibilities? (Open-Ended)

### Civic Engagement

1. How often do you participate in community engagement activities (e.g., voting, providing feedback to local authorities)? (Options: Weekly, Monthly, Rarely, Never)
2. Do you currently face challenges in participating in such activities? (Yes/No)
  - If Yes, what are the main barriers? (Options: Accessibility, Time, Lack of Information, Other)

### During the Testing Period

The ongoing feedback collection during the testing phase will help track:

- **Adoption trends** (e.g., how users adapt to the platform over time).
- **Challenges encountered** while performing specific tasks.
- **Refinement opportunities** before the final deployment.

### Weekly Check-In Questions

These can be embedded as pop-up forms on the platform or sent via email.

### Task-Specific Feedback

1. Were you able to complete your assigned task this week? (Yes/No)
  - If No, what was the issue? (Open-Ended)
2. How easy was it to complete this week's task? (Scale: 1 - Very Difficult, 5 - Very Easy)
3. Did you encounter any errors or issues? If so, please describe.

### Feature Feedback

1. Did you use any accessibility tools (e.g., STT, TTS, font adjustments)? (Yes/No)

- If Yes, how effective were they? (Scale: 1 - Very Ineffective, 5 - Very Effective)
2. Did you explore any new platform features this week? If yes, which ones?

### Engagement and Usability

1. Did the platform help you engage with community activities this week? (Yes/No)
2. How satisfied were you with your overall experience this week? (Scale: 1 - Very Dissatisfied, 5 - Very Satisfied)
3. Any suggestions for improving your experience? (Open-Ended)

### Optional Mid-Testing Focus Group

After 4-6 weeks of testing, consider organising a **mid-testing focus group** to capture in-depth, real-time insights from a subset of users.

#### *Focus Group Discussion Topics*

1. **Adoption Challenges:** What difficulties did you face when first using the platform?
2. **Feature Use:** Which features have you found most valuable so far?
3. **Civic Engagement:** Has the platform encouraged you to participate more in community activities?
4. **Accessibility:** Are there any barriers preventing you from fully utilising the platform?
5. **Suggestions:** What improvements would you recommend at this stage?

### Evaluation Material for Individual User Testing

#### *Online Questionnaire for Users (Post-Session and Exit Survey)*

#### Demographics and Background

1. What is your age?
2. What is your gender?
3. Do you identify with any of the following groups? (Check all that apply: Disability, Elderly, Economically Disadvantaged, General User)
4. Rate your familiarity with using digital platforms before this study. (Scale: 1 - Not Familiar, 5 - Very Familiar)

#### Usability Questions

1. How easy was it to navigate the platform? (Scale: 1 - Very Difficult, 5 - Very Easy)
2. Were you able to complete assigned tasks without external assistance? (Yes/No)
  - If No, please describe the issue.
3. How intuitive did you find the platform's design and features? (Scale: 1 - Not Intuitive, 5 - Very Intuitive)

#### Accessibility Questions

1. Did you use any accessibility features (STT, TTS, font customization, keyboard navigation)? (Yes/No)
2. How satisfied were you with these features? (Scale: 1 - Very Dissatisfied, 5 - Very Satisfied)
3. What additional accessibility features would you suggest? (Open-Ended)

### **Civic Engagement**

1. Did the platform help you feel more engaged in your community? (Scale: 1 - Not at All, 5 - Very Much)
2. Which features best supported your civic engagement efforts? (Open-Ended)
3. How likely are you to use this platform to participate in future civic activities? (Scale: 1 - Very Unlikely, 5 - Very Likely)

### **General Feedback**

1. What did you like most about the platform? (Open-Ended)
2. What would you improve? (Open-Ended)
3. Any additional comments?

## *Focus Group Discussion Guide*

### **Introduction Questions (Icebreaker)**

1. Can you share your initial impressions of the platform?
2. What was the most positive aspect of using it?

### **Main Discussion Topics**

#### **Usability and Accessibility**

- Were the platform's features intuitive and easy to use?
- What challenges did you face when navigating or completing tasks?
- How effective were the accessibility tools (STT, TTS, font customization)?

#### **Civic Engagement**

- Did the platform support your role as a community member or leader?
- How did the platform influence your ability to participate in civic activities?

#### **Features and Feedback**

- Which features did you find most helpful?
- What features do you feel were missing or could be improved?

### **Wrap-Up**

- What are your overall thoughts on the platform?
- Would you recommend it to others? Why or why not?

## Stakeholder Interview Guide

### Introduction and General Feedback

1. What is your role in the community, and how did you interact with the platform?
2. How well do you think the platform aligns with the community's civic engagement needs?

### Usability and Accessibility

1. How easy was it for your organisation to adopt the platform?
2. Were there any accessibility challenges or feedback from the community?

### Impact on Civic Engagement

1. How did the platform facilitate civic participation and decision-making?
2. Were there any observable changes in community involvement during the testing period?

### Future Recommendations

1. What changes or improvements would you suggest to better meet community needs?
2. How do you see the platform scaling to other communities or contexts?

## Platform Analytics

### Key Metrics to Track

1. **Session Length:** Average time users spend per session.
2. **Task Completion Rates:** Percentage of completed tasks (e.g., submitting feedback, voting).
3. **Accessibility Usage:** Number of times STT, TTS, or other accessibility tools were used.
4. **Engagement Metrics:**
  - Number of votes submitted.
  - Number of feedback forms completed.
  - Frequency of logins.
5. **Error Logs:** Frequency and types of errors encountered.
6. **Heatmaps:** Sections of the platform where users spend the most time.

### Notes on Data Processing

All data will be anonymised or pseudonymised during collection and analysis to ensure user privacy. For vulnerable groups, insights will be aggregated and generalised to avoid the identification of specific individuals.

## Annex III: List of Metrics

A complete list of metrics that will be used to evaluate the platform's performance and user engagement across the pilot sites was created and circulated to involved partners for verification. This section includes both technical and user-focused metrics that can be found [here](#).

## Annex IV: Preliminary Testing Scenarios for Platform Functionalities

This annex provides a comprehensive set of preliminary testing scenarios designed to evaluate each identified platform functionality. These scenarios outline the key steps, actions, and expected outcomes necessary for assessing the platform's usability, accessibility, and technical performance. While currently aligned with the functionalities identified during the design phase, these scenarios will be refined and adapted to incorporate the specific Use Case requirements outlined in the main report. The final testing scenarios will be validated and documented in subsequent deliverables (D4.2 and D4.3). The pool of scenarios can be found [here](#).

## Annex V: Template for Cognitive Walkthrough Test Plan for Expert Evaluation

This template provides a structured format for conducting cognitive walkthroughs during the expert testing phase (Phase 0). It includes sections for documenting the testing setup, task evaluations, expert observations, and actionable recommendations.

### Cognitive Walkthrough Test Plan

**Test ID:** [Unique Test Identifier]

**Date:** [DD/MM/YYYY]

**Location:** [Remote/In-Person Testing, Specify Location]

**Observers:** [Names and Roles]

**Experts Involved:** [List of Experts, Roles, and Areas of Expertise]

### Test Setup

**Objective:** [State the purpose of this test, e.g., "To evaluate the usability of the registration process in the platform wireframes."]

### Test Materials:

Wireframe or prototype link/access.

Task instructions for experts.

Evaluation checklist.

**Tasks Evaluated:** [List tasks to be completed by the expert, e.g., registration, navigation, posting content.]

**User Personas:** [Describe the simulated user persona for the expert, e.g., "Visually impaired user," "First-time platform user."]

## Task Evaluation Template

For each task, use the following format:

**Task Name:** [E.g., Registering an Account]

**Scenario:** [Describe the scenario, e.g., "The user is a first-time visitor registering an account to participate in community discussions."]

**Expected Outcome:** [Describe the successful completion of the task, e.g., "The user registers an account and receives a confirmation email."]

### Evaluation:

**Will the user know what to do at this step?** [Yes/No, with explanation.]

**Will the user notice the correct interface elements?** [Yes/No, with explanation.]

**Will the user understand the feedback provided?** [Yes/No, with explanation.]

**Will the user be able to recover from errors?** [Yes/No, with explanation.]

### Expert Observations:

[Detailed notes on issues or challenges encountered, e.g., "The 'Register' button is not prominent enough."]

[Comments on interface clarity, navigation, or accessibility, e.g., "Font size is too small for visually impaired users."]

### Recommended Improvements:

[Actionable suggestions, e.g., "Increase the size and contrast of the 'Register' button."]

### Summary of Findings

**Task Success Rate:** [E.g., "80% of tasks were completed without issues."]

**Critical Issues Identified:** [List key barriers to task completion.]

**Accessibility Observations:** [Notes on STT/TTS functionality, navigation, or compliance with accessibility standards.]

**General Usability Observations:** [E.g., "Navigation was intuitive overall, but experts noted difficulty locating sub-menu items."]

### Recommendations

#### Priority Issues:

[E.g., "Improve visibility of key navigation buttons."]

[E.g., "Ensure STT functionality is available for all input fields."]

**Short-Term Adjustments:** [E.g., "Update button labels to improve clarity."]

**Long-Term Adjustments:** [E.g., "Redesign the onboarding process for new users."]

### Annexes

Annotated screenshots of wireframes highlighting issues.

Recorded observations from the walkthrough session.

Summary tables of expert ratings for each task (if applicable).

## Annex VI: Software Quality testing procedures

The following testing procedures outline the steps to evaluate each software quality metric for the ITHACA platform during Phase 1. These procedures ensure the platform meets the necessary reliability, performance, and security standards under controlled testing conditions.

### 1. Availability and Uptime

- **Objective:** Ensure the platform remains operational for 99.5% of the time, minimising unplanned downtime.
- **Procedure:**
  1. Deploy the platform on test servers with monitoring tools (e.g., Nagios, Zabbix) enabled.
  2. Simulate typical and peak usage conditions over a defined period.
  3. Monitor system uptime continuously and log any downtime incidents.
  4. Analyse downtime frequency and duration using server logs.
- **Tools:** Continuous server monitoring tools (e.g., Zabbix, Pingdom).

### 2. Response Time

- **Objective:** Evaluate how quickly the platform responds to user actions, maintaining response times under 2 seconds.
- **Procedure:**
  1. Simulate user interactions, such as logging in, navigating menus, and submitting forms.
  2. Measure response times for each action using performance testing tools (e.g., JMeter).
  3. Identify any delays and investigate root causes.
  4. Optimise backend processes to improve responsiveness.
- **Tools:** Performance testing tools (e.g., Apache JMeter, Selenium).

### 3. Error Rates

- **Objective:** Minimise system errors to  $\leq 3$  critical errors and  $\leq 10$  minor errors per month.
- **Procedure:**
  1. Monitor error logs during simulated platform usage.
  2. Classify errors as critical (e.g., crashes) or minor (e.g., UI glitches).
  3. Investigate causes of errors and implement fixes.
  4. Re-test to verify resolution and ensure error rates remain within acceptable limits.
- **Tools:** Error log analysis tools (e.g., ELK Stack).

### 4. Stress and Load Capacity

- **Objective:** Ensure the platform can handle 500+ concurrent users and process  $\geq 1,000$  transactions per minute under peak load.

- **Procedure:**
  1. Configure load simulation scenarios using tools like LoadRunner.
  2. Gradually increase the number of simulated users to peak levels.
  3. Measure system performance metrics such as response times, error rates, and data throughput.
  4. Analyse system behaviour under load and identify bottlenecks.
- **Tools:** Load simulation tools (e.g., LoadRunner, Gatling).

## 5. System Downtime

- **Objective:** Minimise unplanned downtime to less than 5 minutes per incident.
- **Procedure:**
  1. Simulate failure scenarios such as server overloads or power interruptions.
  2. Record downtime incidents and duration using downtime monitoring tools.
  3. Analyse system recovery mechanisms and optimise where necessary.
- **Tools:** Downtime monitoring tools (e.g., Pingdom).

## 6. Data Throughput

- **Objective:** Measure the volume of data the platform can process under load, with a target of  $\geq 1,000$  transactions per minute.
- **Procedure:**
  1. Simulate high-volume data processing tasks using load testing tools.
  2. Track transaction rates under varying load levels.
  3. Identify and address any performance bottlenecks.
- **Tools:** Load testing tools (e.g., Apache JMeter, LoadRunner).

## 7. CPU Usage

- **Objective:** Maintain average CPU usage below 70% during peak load.
- **Procedure:**
  1. Monitor CPU usage during stress tests.
  2. Identify high-resource operations and optimise processes.
  3. Ensure adequate server provisioning to handle peak loads.
- **Tools:** System monitoring tools (e.g., Nagios, Zabbix).

## 8. Memory Usage

- **Objective:** Ensure memory usage remains below 75% during peak load.
- **Procedure:**
  1. Monitor memory usage during load simulations.
  2. Identify memory leaks or inefficient memory utilisation.
  3. Optimise memory management in code and adjust server configurations.
- **Tools:** Memory monitoring tools (e.g., New Relic, Nagios).

## 9. Network Latency

- **Objective:** Maintain network latency below 100 ms to ensure fast data communication.

- **Procedure:**
  1. Simulate data transmission scenarios, such as uploading files or retrieving database records.
  2. Measure latency using network monitoring tools.
  3. Optimise network configurations and reduce transmission delays.
- **Tools:** Network monitoring tools (e.g., Wireshark, Nagios).

## 10. Database Query Performance

- **Objective:** Ensure database queries execute within 200 ms.
- **Procedure:**
  1. Simulate typical database operations such as retrieving user data or submitting feedback.
  2. Measure query execution times using database performance testing tools.
  3. Optimise database indexing and query design for faster execution.
- **Tools:** Database performance tools (e.g., Query Monitor).

## 11. Security Vulnerabilities

- **Objective:** Achieve zero critical security vulnerabilities.
- **Procedure:**
  1. Perform penetration testing using tools like OWASP ZAP.
  2. Identify vulnerabilities in authentication, data transmission, and storage.
  3. Implement necessary security measures and verify fixes through re-testing.
- **Tools:** Penetration testing tools (e.g., OWASP ZAP, Burp Suite).

## 12. Data Integrity

- **Objective:** Ensure 100% data accuracy and consistency.
- **Procedure:**
  1. Test data input, processing, and output for consistency.
  2. Validate the accuracy of user-submitted data and system-generated reports.
  3. Use automated integrity testing to identify discrepancies.
- **Tools:** Automated integrity testing tools (e.g., DbFit).

## 13. Backup and Recovery Time

- **Objective:** Achieve full backup and recovery within one hour.
- **Procedure:**
  1. Perform scheduled backups and measure backup duration.
  2. Simulate recovery scenarios and measure restoration times.
  3. Optimise backup and recovery procedures as necessary.
- **Tools:** Backup and recovery tools (e.g., Veeam, Acronis).

## Procedure for Component (Unit) Testing

Component (unit) testing is a critical step in the validation process, focusing on the verification of individual platform modules to ensure they function as intended before being integrated into the

larger system. This process isolates specific components—such as the voting module, AI-driven features, or accessibility tools (STT; TTS)—to identify and address any defects early in the development lifecycle. The testing procedure involves structured test cases designed to evaluate both expected and edge-case scenarios, ensuring robustness and reliability.

## 1. Preparation Phase

**Identify Components for Testing:** Key components to be tested include user authentication modules, voting and polling mechanisms, AI features for content summarization and explainability, and accessibility tools like STT and TTS.

**Develop Test Cases:** For each component, test cases are designed to cover:

- Expected scenarios (e.g., a user successfully logging in).
- Edge cases (e.g., invalid credentials or input errors).
- Failure scenarios (e.g., system handling of missing data or network interruptions).

**Set Up Testing Environment:** A controlled testing environment is established with mock data inputs, ensuring the component can be tested independently of other system parts.

## 2. Testing Phase

### Execution of Test Cases:

- Execute the prepared test cases for each component.
- For example:
  - **Voting Module:** Test voting functionality with varying inputs, such as single-choice, multiple-choice, and custom-ranking polls, to ensure accurate submission and data storage.
  - **AI Features:** Test AI-driven content summarization for accuracy and relevance, ensuring it produces comprehensible and concise outputs.
  - **STT and TTS Tools:** Validate STT's ability to accurately transcribe various speech patterns and TTS's capability to read platform content clearly.
- Record any anomalies or deviations from expected results.

### Boundary Testing:

- Test components against boundary conditions, such as character limits in text fields or maximum concurrent users for a specific module.

### Error Handling and Recovery:

- Evaluate the component's ability to handle errors gracefully, such as invalid user inputs or server unavailability.
- Test recovery mechanisms, ensuring the component returns to a stable state after an error.

### Performance Testing:

- Assess response times and resource utilisation for each component to ensure they meet performance criteria.
- For example, verify that the AI-driven summarization completes within a specified time limit.

### 3. Integration Readiness Check

#### Component Interaction Simulation:

- Simulate basic interactions between the tested component and a limited set of related modules.
- For instance, test the voting module's integration with the database for recording and retrieving poll results.

#### Verification of Outputs:

- Ensure that the outputs of each component are accurate and formatted correctly to be utilised by other system modules during integration.

### 4. Reporting and Iteration

#### Document Results:

- Record test outcomes, including details of any failures, observed behaviour, and potential causes.
- Maintain a log of all changes made to the component during testing.

#### Refinement and Retesting:

- Address any identified issues, refine the component as necessary, and repeat the test cases to verify fixes.

#### Sign-Off for Integration:

- Once a component has passed all test cases, it is marked as ready for integration into the larger system.

The component (unit) testing phase ensures that individual modules, such as the voting system and AI features, function as intended in isolation. By addressing issues early, this approach reduces the complexity of troubleshooting during integration testing and improves the overall reliability of the platform. Each component will enter the integration phase with proven functionality, laying the foundation for a stable and cohesive system.

## Procedure for Integrated Platform Testing

Integrated platform testing focuses on evaluating the seamless interaction between different components of the ITHACA platform. This phase ensures that all modules work cohesively, verifying data flow, interface compatibility, and system stability. Additionally, security and stress testing

procedures are integrated to assess the platform's resilience under high loads and ensure compliance with data protection regulations.

## 1. Preparation Phase

**Define Testing Scenarios:** Develop scenarios that involve multiple components, such as:

- The integration of accessibility tools (e.g., STT/TTS) with content management systems.
- Database queries handling user inputs, such as voting results or feedback submissions.
- AI-driven summarization interacting with user-provided data and presenting outputs.

**Set Up Testing Environment:**

- Deploy the full platform in a controlled environment that mimics real-world conditions.
- Ensure all modules, APIs, and databases are connected and operational.

**Security Compliance Checklist:**

- Review GDPR and other relevant data protection regulations.
- Identify critical areas for security testing, such as data storage, transmission, and user authentication.

## 2. Testing Phase

### Functional Integration Testing

**Data Flow Verification:**

- Test end-to-end data flow between modules, such as:
  - User input flowing into the database and being retrieved for visualisation or reports.
  - Accessibility tools interacting with input fields and presenting outputs accurately.
- Verify the correctness and completeness of data transfer between components.

**Interface Compatibility:**

- Ensure UI components align with backend processes, such as AI-driven tools providing accurate and timely outputs on the front end.
- Test the responsiveness of various interface elements during simultaneous interactions.

**Cross-Module Functionality:**

- Simulate complex workflows involving multiple components, such as a user submitting feedback with STT, which is stored in the database and then summarised by the AI module.

## Security Testing

### Vulnerability Assessment:

- Use penetration testing tools (e.g., OWASP ZAP, Burp Suite) to identify security vulnerabilities in the platform, such as injection flaws, broken authentication, and misconfigurations.

### Data Protection Compliance:

- Evaluate data handling processes to ensure compliance with GDPR and other regulations:
  - Encrypt sensitive user data during storage and transmission.
  - Test anonymization processes for feedback and voting data.
- Verify user authentication and access control mechanisms, ensuring restricted access to sensitive data.

### Incident Simulation:

- Simulate potential security breaches, such as unauthorised access attempts, and evaluate the platform's response mechanisms.

## Stress Testing

### 1. Simulate Peak Loads:

- Gradually increase the number of concurrent users to test system performance under heavy load.
- Use load simulation tools (e.g., LoadRunner, Gatling) to create scenarios with 500+ users performing simultaneous actions, such as voting, commenting, or uploading files.

### 2. Measure Performance Metrics:

- Evaluate response times, data throughput, CPU/memory usage, and network latency under stress conditions.
- Identify any bottlenecks or performance degradation during high-traffic scenarios.

### 3. Recovery and Resilience:

- Simulate unexpected failures, such as server crashes or database overloads, and assess recovery time and data integrity post-failure.

## 3. Analysis and Iteration

### 1. Documentation of Results:

- Record all test outcomes, including:
  - Functional errors during integration.
  - Security vulnerabilities and their severity.
  - Stress test results, including performance metrics and failure points.

## 2. Issue Resolution:

- Address identified issues, optimise data flow, and enhance interface compatibility.
- Implement fixes for security vulnerabilities and retest to confirm resolution.

## 3. Iteration:

- Repeat integration tests with resolved components to ensure seamless operation.

Integrated platform testing ensures that all modules interact effectively, supporting the platform's intended functionalities. It verifies data flow, interface compatibility, and security compliance, while stress testing evaluates the platform's resilience under high loads. This phase ensures the platform is robust, secure, and ready for deployment in real-world conditions.

## Procedure for Security and Stress Testing

This section outlines the procedures for conducting **security testing** to ensure platform resilience against cyber threats and compliance with data protection regulations, as well as **stress testing** to evaluate system performance under peak load conditions.

### 1. Security Testing

**Objective:** Identify vulnerabilities, ensure compliance with data protection regulations (e.g., GDPR), and validate the platform's ability to withstand potential cyber threats.

#### Testing Steps:

##### 1. Vulnerability Assessment:

- Conduct an automated vulnerability scan using tools such as OWASP ZAP or Burp Suite to identify potential threats, including:
  - Injection flaws (e.g., SQL Injection).
  - Broken authentication.
  - Security misconfigurations.
- Analyse scan results to prioritise high-risk vulnerabilities.

##### 2. Penetration Testing:

- Simulate real-world cyberattacks to test the platform's defence mechanisms.
- Perform the following tests:
  - **Authentication Security:** Validate password policies, session management, and multi-factor authentication mechanisms.
  - **Data Transmission Security:** Test encryption of sensitive data during transfer using HTTPS protocols.
  - **Access Control Testing:** Verify that unauthorised users cannot access restricted features or data.

##### 3. Data Protection Compliance:

- Assess the platform's adherence to GDPR by evaluating the following:
  - Data anonymization for sensitive information.
  - Consent mechanisms for data collection and usage.
  - Secure storage and processing of personal data.
- Test the implementation of **data encryption protocols** for stored data and ensure that all sensitive information is encrypted by default.

##### 4. Incident Response Simulation:

- Simulate security breach scenarios (e.g., unauthorised access attempts or DoS attacks) to test the platform's detection and mitigation mechanisms.
  - Validate logging and alerting systems for incident tracking.
- 5. Validation and Retesting:**
- Address identified vulnerabilities and perform retesting to ensure fixes are effective.
  - Ensure **zero critical security vulnerabilities** remain post-testing.

**Tools:**

- OWASP ZAP, Burp Suite, Nessus for vulnerability and penetration testing.
- GDPR compliance tools for data protection validation.

## 2. Stress Testing

**Objective:** Evaluate the platform's capacity to handle heavy user loads, ensuring performance remains stable and responsive during peak conditions.

**Testing Steps:**

- 1. Scenario Definition:**
  - Define realistic peak load scenarios, such as:
    - 500+ concurrent users submitting feedback or voting simultaneously.
    - Processing 1,000+ transactions per minute during a community event.
  - Identify critical platform features to monitor, including response time, data throughput, and error rates.
- 2. Load Simulation:**
  - Use tools such as **LoadRunner** or **Gatling** to simulate user interactions under increasing loads.
  - Gradually ramp up the number of simulated users to peak levels, monitoring system behaviour throughout the test.
- 3. Performance Metric Monitoring:**
  - Measure key metrics during the test:
    - **Response Time:** Ensure user actions are processed within 2 seconds.
    - **System Uptime:** Maintain 99.5% availability.
    - **Error Rates:** Limit to ≤ 3 critical errors and ≤ 10 minor errors per month.
    - **Data Throughput:** Achieve ≥ 1,000 transactions per minute.
  - Track resource utilisation (CPU, memory) and network latency to identify bottlenecks.
- 4. Failure Simulation:**
  - Simulate failure scenarios, such as sudden server overloads or network outages, to evaluate system resilience.
  - Assess recovery time and system stability post-failure.
- 5. Analysis and Optimization:**
  - Analyse test results to identify performance bottlenecks, such as high CPU usage or memory leaks.
  - Optimise resource allocation, such as database queries or load balancing configurations, to improve system performance.
- 6. Validation and Retesting:**

- Retest after optimization to ensure the platform meets performance targets under stress conditions.

### Tools:

- LoadRunner, Apache JMeter, Gatling for load simulation.
- Nagios, Zabbix for monitoring CPU, memory, and network latency.
- **Security Testing:**
  - Identification and resolution of all critical security vulnerabilities.
  - Full compliance with data protection regulations, including GDPR.
  - Verified resilience against common cyber threats, ensuring user trust.
- **Stress Testing:**
  - Validation of the platform's ability to handle 500+ concurrent users and process  $\geq$  1,000 transactions per minute.
  - Stability and responsiveness maintained under peak load conditions.
  - Identification of performance bottlenecks and successful implementation of optimizations.

These procedures ensure the ITHACA platform is both secure and capable of delivering high performance during real-world usage scenarios.

## Annex VII: User acceptance and experience testing (Phase 1 & 2)

### Procedure for Each User in Phase I: Controlled Setting Testing

Phase I testing of the ITHACA platform will be conducted in a controlled environment to ensure a systematic and reliable evaluation of the platform's usability, performance, and accessibility. The procedure for each user will follow a structured process, ensuring consistency and comprehensive data collection.

#### 1. Introduction and Onboarding

##### Welcome and Introduction:

- Each user will be greeted and introduced to the purpose of the testing session.
- A brief overview of the ITHACA platform's objectives and features will be provided, tailored to the user's level of familiarity with digital tools.

##### Consent and Demographics:

- Users will sign a consent form agreeing to participate in the testing and allowing their data to be anonymized for analysis.
- Basic demographic data (e.g., age, digital literacy level, any disabilities) will be collected to ensure diversity and inclusivity.

##### Training:

- Users will receive a short tutorial on how to navigate the platform.
- Accessibility features such as STT, TTS, and font customization will be demonstrated, ensuring users understand how to use them if needed.

## 2. Testing Tasks

Each user will complete a predefined set of tasks that represent typical interactions with the platform. Tasks will be standardised to ensure comparability across participants. The following are mere suggestions as the platform is not yet available.

### **Task 1: Account Setup and Login:**

- Users will create a new account and log into the platform.
- They will be instructed to customise their profiles, including selecting accessibility options if needed.

### **Task 2: Navigation and Information Retrieval:**

- Users will locate specific sections of the platform, such as "Community Discussions" or "Voting."
- They will navigate to a designated post and review its content.

### **Task 3: Voting and Poll Interaction:**

- Users will participate in a mock voting session by selecting options in a poll.
- They will interact with voting features, including STT input for votes where applicable.

### **Task 4: Content Contribution:**

- Users will create a new post using the platform's content management tools.
- They will input free text, upload a file, and enable accessibility options like TTS for their content.

### **Task 5: Providing Feedback:**

- Users will complete a feedback form regarding their experience using the platform.
- They will explore additional feedback submission options, such as reporting an issue or suggesting a feature.

## 3. Observations and Data Collection

### **Facilitated Monitoring:**

- A facilitator will observe each user's interaction with the platform, noting areas where the user encounters difficulties, hesitations, or errors.
- Observations will focus on task completion time, number of errors, and any assistance required.

### **Automated Data Logging:**

- The platform will track metrics such as click paths, time spent on tasks, and task completion rates.
- Analytics tools will record performance data like response times and system stability during user interactions.

**User Feedback Collection:**

- Users will complete a post-test survey or usability questionnaire (e.g., SUS or IsoMetrics).
- Open-ended questions will capture qualitative insights into their experience, such as suggestions for improvement or perceived challenges.

#### *4. Accessibility and Emotional Response Evaluation*

**Accessibility Tools Testing:**

- Users with disabilities will be specifically tasked with testing STT, TTS, and other accessibility features.
- Feedback will be collected to assess satisfaction and usability of these tools.

**Emotional Impact Assessment:**

- Users will be asked to rate their emotional responses (e.g., frustration, empowerment, or satisfaction) during the session.
- Facilitators may conduct brief interviews to explore these emotional responses further.

#### *5. Exit and Follow-Up*

**Debriefing:**

- Users will be thanked for their participation and provided with a summary of the testing process.
- They will have the opportunity to ask questions or provide additional verbal feedback.

**Follow-Up Communication:**

- Users will be informed about the outcomes of the testing phase and how their input will be utilised to refine the platform.
- If applicable, users may be invited to participate in future testing phases or focus groups.
- Task completion rates and times, revealing usability challenges.
- User satisfaction scores and qualitative feedback to guide platform refinement.
- Insights into the effectiveness of accessibility tools and their alignment with user needs.
- System performance data, identifying potential technical bottlenecks.

## Detailed Procedure for User Experience Testing in Real-Life Settings (Phase 2)

This section outlines step-by-step instructions for conducting the user experience testing in real-life settings, including procedures for individual user testing, focus groups, and stakeholder interviews. Each step is broken down to ensure clarity and replicability, enabling facilitators with minimal prior knowledge to carry out the testing effectively.

### Procedure for Individual User Testing

#### *Preparation*

##### 1. Recruitment:

- Identify 20 participants per pilot site. The target composition is:
  - 5 individuals with disabilities (e.g., visually impaired, hearing impaired, or motor impairments).
  - 5 older users aged 65 or older.
  - 5 individuals from economically disadvantaged backgrounds with limited digital literacy.
  - 5 general users, including active community members or civic leaders.
- Use recruitment channels such as local NGOs, municipal outreach programs, and social media advertisements to ensure diversity.
- Confirm participant consent through signed or digital consent forms provided in the user's preferred language.

##### 2. Technical Setup:

- Ensure all participants have access to devices that support the platform (e.g., desktop, tablet, or smartphone).
- Prepare user accounts for participants, including pre-configured accessibility settings for users with disabilities.
- Provide participants with a unique login ID and password. Include instructions for logging in via email or printed guides.

##### 3. Training:

- Conduct a 15–20-minute onboarding session, either in person or via video call, to demonstrate:
  - Logging into the platform.
  - Key features such as submitting feedback, participating in polls, and using STT/TTS.
  - Accessibility tools (e.g., font customization, navigation options).
- Provide printed or digital user manuals for reference.

#### *Testing Procedure*

##### 1. Day 1: Initial Setup and Orientation:

- Schedule a call with each participant to ensure successful login and navigation through the platform.
- Have participants complete a brief introductory survey via the platform to assess initial impressions of usability.

##### 2. Usage Period:

- Participants will use the platform over two tasks. Couple-of-days tasks will be assigned, such as:
  - **Day 1–2:** Explore the homepage, review civic updates, and participate in a discussion forum.
  - **Day 3–4:** Submit a vote on a community project and provide feedback via the feedback form.
  - **Day 5–6:** Use STT to complete a form and TTS to review content (for users required to use them).
  - **Day 7–8:** Test accessibility features, including font adjustments and keyboard navigation (for users required to use them).
  - **Days 9–14:** Perform self-directed tasks, such as browsing updates, interacting with the guidance bot.
- 3. Data Collection:**
  - **During Use:**
    - Track usage analytics (e.g., session length, number of interactions, task completion rates).
    - Capture system logs for errors or issues encountered.
  - **Post-Session Surveys:**
    - At the end of each week, participants will complete a survey embedded within the platform to assess their experience with the tasks performed.
- 4. Exit Survey:**
  - After two-week testing period, participants will complete a comprehensive exit survey evaluating their overall experience, satisfaction, and the platform's contribution to their civic engagement.

## Procedure for Focus Groups

### *Preparation*

- 1. Participant Selection:**
  - Invite 5–7 participants from the user testing group for each focus group session.
  - Ensure representation from vulnerable groups, with at least:
    - 2 individuals with disabilities.
    - 2 older users.
    - 1 economically disadvantaged user.
    - Additional general users as needed.
- 2. Facilitation Setup:**
  - Select a neutral and accessible venue for in-person sessions or a reliable video conferencing tool for virtual sessions.
  - Prepare recording equipment (e.g., audio recorders) and note-taking tools.
  - Have a trained moderator with knowledge of the platform and civic engagement facilitate the session.

### *Focus Group Session*

- 1. Introduction (15 minutes):**
  - Moderator introduces themselves and outlines the purpose of the session.
  - Provide an overview of the platform and testing objectives.

- Reiterate confidentiality and obtain verbal consent to record the session.
- 2. **Discussion Topics** (60 minutes):
  - **Usability:** What aspects of the platform were intuitive? Which features were difficult to use?
  - **Accessibility:** Were accessibility features sufficient? What improvements are needed?
  - **Civic Engagement:** Did the platform help you participate in community activities? How could it better support civic roles?
  - **General Feedback:** What features were most useful? What additional functionalities would you recommend?
- 3. **Closing** (15 minutes):
  - Moderator summarises key points raised during the discussion.
  - Participants are thanked for their contributions and given a small token of appreciation (e.g., gift card or certificate).

**Output:** Transcribe recordings and analyse themes using qualitative analysis tools like NVivo or manual coding.

## Procedure for Stakeholder Interviews

### *Preparation*

1. **Participant Selection:**
  - Identify 5 stakeholders per pilot site, including:
    - 2 municipal representatives.
    - 2 community leader.
    - 1 NGO representative.
  - Send an invitation with an explanation of the interview purpose and a proposed schedule.
2. **Interview Setup:**
  - Conduct interviews in person or virtually, depending on stakeholder preference.
  - Prepare a semi-structured interview guide tailored to their role and interactions with the platform.

### *Interview Process*

1. **Introduction** (5 minutes):
  - Interviewer introduces themselves and provides an overview of the session's objectives.
  - Confirm participant consent for recording and explain confidentiality.
2. **Interview Questions** (30–40 minutes):
  - **Platform Adoption:** How did you use the platform within your role? Were any challenges encountered during adoption?
  - **Impact Assessment:** Did the platform facilitate improved community engagement or decision-making processes?
  - **Feedback on Features:** Which platform features were most beneficial? Were there any missing functionalities?
  - **Future Recommendations:** How could the platform be scaled or improved for broader adoption?

### 3. Closing (5 minutes):

- Thank the participant for their insights.
- Summarise the key points discussed and explain how their feedback will inform further development.

**Output:** Transcribe interviews and summarise insights in a report.

## Annex VIII: Structure of the ITHACA Platform User Manual

### 1. Introduction

- **1.1 Purpose of the User Manual**

Explanation of the manual's purpose, audience, and scope.

- **1.2 Overview of the ITHACA Platform**

Brief description of the platform's features, goals, and intended use cases.

### 2. Getting Started

- **2.1 System Requirements**

Hardware, software, and internet requirements for accessing the platform.

- **2.2 User Registration and Account Setup**

Step-by-step guide for creating an account and logging in.

- **2.3 Navigating the Platform**

Overview of the main interface, menus, and key features.

### 3. Features and Functionalities

- **3.1 Participatory Tools**

Explanation of voting, polls, surveys, and feedback submission.

- **3.2 Accessibility Features**

Use of accessibility tools like Text-to-Speech (TTS) and Speech-to-Text (STT).

- **3.3 Notifications and Updates**

How to manage notifications and follow project updates.

- **3.4 Personalization and Settings**

Customization options, such as language preferences and user settings.

### 4. Use Case Scenarios

- **4.1 Civic Poll Participation**

Example scenarios demonstrating how to engage in civic polls.

- **4.2 Event Discovery and Feedback**

Instructions for browsing events, registering, and providing feedback.

- **4.3 AI-Driven Recommendations**

How to interpret and interact with personalized recommendations.

## 5. Troubleshooting and Support

- **5.1 Common Issues and Solutions**

FAQs and solutions to common technical and usability issues.

- **5.2 Contacting Technical Support**

How to report issues and seek assistance.

## 6. Data Privacy and User Security

- **6.1 Data Privacy Policies**

Overview of data handling and GDPR compliance.

- **6.2 User Security Tips**

Recommendations for secure account management and online interactions.

## 7. Testing and Feedback

- **7.1 Instructions for Test Users**

Guidance on participating in Phase 1 and Phase 2 tests.

- **7.2 Feedback Submission**

How to provide structured feedback during tests.

- **7.3 Reporting Usability Issues**

Logging issues and suggestions for improvement.

## 8. Appendices

- **8.1 Glossary of Terms**

Definitions of technical terms and abbreviations used in the manual.

- **8.2 Sample Scenarios**

Detailed examples of user interactions with key features.

- **8.3 Contact Details**

List of relevant contacts for technical and administrative support.