# Improving Public Procurement for Digital Work Environments Through Research-Practitioner Collaboration

Marta Larusdottir[1][0000-0003-0256-5710] and Åsa Cajander[2][0000-0001-7472-2215]
[1]Reykjavik University, Reykjavik, Iceland
[2]Uppsala University, Uppsala, Sweden
marta@ru.is

**Abstract.** Public authorities face significant challenges when procuring software systems that effectively support their operations and employees. The mandatory Request for Proposal (RFP) process often prioritises cost and technical specifications, with limited attention to usability and digital work environment factors. This paper addresses this gap by exploring how usability evaluations and performance criteria can be systematically integrated into the tendering process for public sector procurement. Drawing on collaboration between academic researchers and practitioners, the study examines two case studies: (1) the selection of an agile development team through RFP for developing a financial support application and (2) the selection of a running collaboration and information management system (CIMS) through RFP for government ministries. In the first case, team collaboration, user experience focus, and code quality were evaluated as performance factors in the selection criteria in parallel with the estimated cost of developing the system. In the second case, the selection factors were cost and quality, including usability metrics such as effectiveness, efficiency, and satisfaction. These cases illustrate how research-practitioner collaboration can bridge the gap between academic principles and professional practices, advancing the design and procurement of software that extends the quality of the digital work environment. The findings contribute actionable insights into enhancing public sector RFPs, offering a replicable model for integrating usability and performance evaluation frameworks to improve digital work environments.

Keywords: Digital Work Environment, Request for Proposals, Public Authorities, Selection Criteria, Usability Evaluation.

## 1. Introduction

IT systems' usability and user experience (UX) are critical for workplace efficiency, employee satisfaction, and overall organisational success [1]. However, many workplace systems suffer from usability flaws, leading to inefficiencies, frustration, and burnout [2–4]. Despite increasing recognition of usability, it is often deprioritised during system development, with user-centred design practices implemented inconsistently [5,6].

When public authorities acquire or develop new software systems to support their operations, they must adhere to a complex and regulated selection process. European

Union legislation mandates that this process be transparent and competitive, requiring public authorities to issue a Request for Proposal (RFP). These RFPs outline the requirements, selection criteria, and expectations for custom software development and the procurement of off-the-shelf solutions. Despite the pivotal role software plays in improving the efficiency and effectiveness of public services, the criteria in these processes often prioritise cost and technical specifications over usability and other factors critical to the digital work environment [7–9].

In practice, RFPs typically include two key sections: (1) the system requirements and needs and (2) the selection criteria [8]. Commonly, the selection process places significant emphasis on cost, often awarding contracts to the lowest bidder, even at the expense of quality [9]. For example, a study of software companies in Denmark found that developers primarily focused on what public authorities explicitly required, with little attention to quality factors such as user experience (UX) or security, unless explicitly stated in the RFP [10].

While some RFPs include methods for evaluating software development methods, such as user testing or defined performance criteria, usability considerations are rarely integrated as a core element [10]. Furthermore, selection criteria infrequently assess the development team's competencies, such as their experience with user-centred design, which could significantly contribute to the quality of the final product. Studies have shown that including usability-related requirements can positively influence the project outcome, leading to systems that are functional but also reliable, maintainable, and usable [11].

Software procurement is a crucial activity that has the potential to improve software users' work environments. Additionally, many studies have been conducted on the benefits of conducting usability evaluations for improving the usage of software systems. While previous research has acknowledged the importance of including usability in public sector software procurement [8,9] and thereby improving the work environment, only a small number of research studies have systematically integrated usability evaluation criteria into real-world procurement processes. Some usability methods, like the SUS questionnaire [24,25], have been validated for years and have been reliable, but are rarely used in software procurement. This study builds on a collaborative effort between academic researchers and practitioners in the public sector to address this gap.

This study builds on existing work by examining two empirical case studies where usability-focused evaluation methods were applied in public tenders. By comparing workshop-based selection with structured usability testing, we provide new insights into the effectiveness of these approaches in ensuring user-centred procurement decisions. These findings contribute a replicable model for integrating usability and performance evaluation frameworks into public procurement, which can inform future policy and practice. Combining theoretical insights with real-world applications examines how usability and performance criteria can be effectively integrated into RFPs to improve software procurement outcomes. The findings demonstrate actionable ways to incorporate usability-focused evaluation frameworks into the tendering process, ensuring that public sector software systems prioritise user-centred outcomes.

Through two case studies, the paper explores (1) the selection of an agile development team for a financial support application and (2) the usability evaluation of collaboration and information management systems (CIMS) for government ministries.

In both cases, researchers and practitioners collaborated to design and implement evaluation methods, fostering a bridge between academic principles of interaction design and professional practices in the public sector.

This paper aligns with the ARPPID conference theme by addressing the critical need for cooperation between academia and industry in interaction design. The main research question guiding this study is:

*"How can usability evaluations and performance criteria, developed through collaboration between academic researchers and practitioners, improve the Request for Proposal process for public authorities?"*

## 2. Background

In this section, we give an overview of the background literature related to digital work environments, usability evaluations and usability evaluations as a part of request for proposals in tendering processes.

### 2.1 Digital Work Environments

The digital work environment refers to the interconnected system of technologies, processes, and organisational factors that shape how employees interact with digital tools in their daily tasks. As defined in the report on Artificial Intelligence and Digitalisation [12], "the digital work environment encompasses the physical, organisational, social, and cognitive problems and opportunities resulting from the digitalisation of work support systems and tools." This comprehensive definition highlights the technical aspects and the broader organisational and individual impacts, emphasising the need for a holistic approach to understanding and improving digital work environments.

Virpi Roto's work extends this understanding by introducing concepts such as "designing for work engagement", which advocates for systems that support users' emotional and cognitive involvement. Roto et al. [13] argue that usability must evolve beyond essential task support to foster meaningful interactions that motivate and engage users in their work environments.

Using the Health-Technology-Organisation (HTO) concept and contributions from HCI research, the digital work environment can be analysed across three primary dimensions [14]:

- Physical and Ergonomic Factors: These include hardware design and workspace configurations. Poor ergonomics can lead to physical discomfort and decreased productivity, while optimised setups improve well-being and efficiency.
- Cognitive Usability: Interfaces must align with users' mental models to ensure intuitive interactions. High cognitive load from poorly designed systems leads to frustration, errors, and diminished productivity.
- Organisational and Social Aspects: Digital tools shape team communication and decision-making. Systems designed with collaboration in mind foster better outcomes and a more substantial workplace culture.

Despite advancements, challenges persist in creating digital work environments that effectively balance usability and organisational needs. Forsman and Eriksson (2023) highlight the risks of technostress and resistance when systems are misaligned with user workflows.

### 2.2 Usability Evaluations

Usability evaluation is a cornerstone of Human-Computer Interaction (HCI), ensuring that systems are functional, user-centred, and satisfying to use. According to the ISO 9241-210 standard [15], usability is "the extent to which a system, product, or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use." Evaluating usability provides insights into user behaviour and identifies design flaws, enabling systems to better align with user needs.

Early methods, such as heuristic evaluations introduced by Nielsen and Molich [16], remain foundational in HCI, offering a cost-effective way to identify usability issues. Think-aloud protocols and cognitive walkthroughs are also widely used, allowing researchers to uncover task-specific challenges and cognitive bottlenecks during interaction [17]. More recently, advancements such as automated usability testing, remote platforms, and AI-driven analysis have enabled more scalable and precise evaluations [18].

Despite their importance, usability evaluations are often underutilised in complex socio-technical systems like public sector IT. Larusdottir et al.[19,19–22] highlighted that usability feedback in large projects often faces integration challenges. Poor usability can lead to inefficiencies, frustration, and resistance, making systematic evaluation essential.

### 2.3 Usability Evaluations in Requests for Proposals

Integrating usability evaluation into the Requests for Proposals (RFP) process is critical for ensuring that procured IT systems meet user needs and align with organisational goals and thereby improves the digital work environment of the employees. In public sector procurement, RFPs typically prioritise cost and technical specifications, often at the expense of usability. However, as Jokela et al. [8] argue, embedding usability considerations into RFPs can significantly enhance the quality and effectiveness of procured systems by emphasising user-centred outcomes.

Usability evaluation in tendering can take several forms, such as defining usability criteria in the RFP, requiring vendors to demonstrate user-centred design processes, or incorporating usability testing as part of the evaluation process. For instance, Tarkkanen and Harkke [10] highlight the importance of specifying performance metrics, such as task completion rates or user satisfaction scores, to ensure that usability is objectively assessed during vendor selection. These practices help ensure that systems are evaluated not only for their functional capabilities but also for their impact on the user experience.

Despite its benefits, integrating usability evaluation into the tendering process faces several challenges. Larusdottir et al. [9] note that usability requirements are often poorly defined in RFPs, leading vendors to deprioritise usability during development.

Additionally, usability evaluation may require additional time and resources, which can conflict with tight procurement timelines. However, researchers argue that the upfront investment in usability evaluation can result in long-term savings by reducing errors, inefficiencies, and user frustration [10].

Public sector organisations can ensure that procured systems are cost-effective, functional and user-centred by incorporating usability evaluation into Requests for Proposals. For HCI researchers, this represents an opportunity to explore methods for standardising usability requirements and developing scalable evaluation frameworks that align with the unique constraints of public procurement.

## 3. Method

This study utilised Rapid Qualitative Analysis (RQA) to explore processes and outcomes related to the Requests for Proposals (RFP) in public procurement, focusing on integrating work environment considerations within case contexts. RQA was chosen because it emphasises iterative and collaborative discussions to derive meaningful insights [23]. The RQA fitted well for the analysis in the paper, since the two authors analysed and discussed the findings in collaboration. The analysis centred on two case studies: selecting an agile development team for a financial support application and evaluating collaboration and document management systems. The cases were analysed in the following categories: Purpose (the purpose of using this method); Assessment focus (the focus of the assessment within the method); Strength (the strength of using the method); Challenges (challenges of using the method); and Best use case (when the method will be of best use).

Data were derived from workshop observations, case documentation, and reflective discussions among the research team. Instead of formal coding, the analysis relied on iterative discussions during the collaborative writing process. Collaborative discussions during the writing process were used to refine the analysis, ensuring that findings were grounded in the empirical data and aligned with theoretical perspectives on public procurement and system development. This comparative approach allowed the research to identify practical insights while maintaining a connection to broader academic frameworks.

## 4. The Cases

In this section, we will describe the two cases, by describing the systems being developed or bought, the selection process, and the results from the selection process.

### 4.1 Case 1: Selecting a team for developing a financial support application

This case examines the selection of a development team for developing a financial support application commissioned by a public authority in Northern Europe. The selection was done by defining an RFP, in which development companies were able to participate. This case was initially detailed in a prior study, which focused on selecting the best agile development team to meet these objectives [26]. The current discussion

extends the analysis by particularly the process of evaluating the digital work environment aspects of the selection process.

#### 4.1.1 The financial support system

The project aimed to create a user-centred digital system that streamlined the application process for financial assistance while ensuring accessibility, operational efficiency, and regulatory compliance. The system targeted three main user groups: citizens, administrative employees, and auditors. Citizens, including individuals with intellectual disabilities, required an accessible and intuitive interface to submit applications. Administrative employees needed a reliable system to manage and process submissions efficiently, while auditors were tasked with reviewing application logs for compliance with financial regulations.

#### 4.1.2 Evaluation process

The development team was selected through a workshop-based evaluation process. Five teams were pre-qualified based on technical requirements outlined in the RTF and invited to participate in a one-day workshop simulating real-world collaboration and development scenarios, teams A to E. The teams got 4 predefined user stories, to work on during the workshop. The teams could decide, if they worked on all of them, or narrowed their project down to working on one or two stories. As an example, one of the user stories was: "*As a citizen of Reykjavik city that has impaired intellectual ability, I want to be able to apply for financial assistance via web/mobile so that I can apply in a simple and easy-to-understand manner*". The other three user stories were directed to the employees supporting the applications for financial support.

Each workshop began with a planning meeting, where teams demonstrated their ability to plan and prioritise tasks. During the development session, teams worked on the predefined user stories designed to assess their technical capabilities and focus on usability. The process concluded with a presentation from the team after the one-day workshop, including presentations of their deliverables and their relevance to the identified user groups' needs.

This approach provided a practical mechanism to assess how teams balanced technical execution with user-centred practices and collaborative problem-solving.

The teams' performance and deliverables were evaluated on three primary performance indicators. First, team collaboration and user-centred focus were assessed by observing their approach during the planning meeting and their presentation at the end of the workshop. Second, the quality and completeness of predefined user stories, such as creating accessible interfaces and implementing features for audit compliance, were analysed by two experts in those areas. Finally, code quality was evaluated based on accessibility, maintainability, and security considerations. These data points provided a comprehensive view of the teams' capabilities and suitability for the project.

#### 4.1.3 Results from the evaluation process

The workshop revealed notable differences in team performance. The selected team excelled in understanding user needs, particularly accessibility-related ones, and demonstrated an ability to incorporate this understanding into their design process.

Their deliverables reflected a strong commitment to usability and technical excellence, and effective collaboration during the workshop. Other teams, while technically competent, lacked a clear focus on user-centred design, which affected their overall evaluation. The results can be seen in Table 1.

**Table 1.** The total points that each team received for the three performance factors evaluated

| Performance factor | Team A | Team B | Team C | Team D | Team E |
|---|---|---|---|---|---|
| Team collaboration and UCD focus<br>*max 25 points* | **25,0** | 12,4 | 9,4 | 7,6 | 19,4 |
| Delivery of user stories<br>*max 10 points* | 3,3 | 6,7 | 6,7 | **10,0** | 6,7 |
| Quality of code<br>*max 35 points* | 22,0 | 16,2 | 18,0 | **22,4** | **22,4** |
| **Total**<br>***max 70 points*** | **50,3** | 35,3 | 34,1 | 40,0 | 48,5 |

The procurement process shifted towards user-centred outcomes in public sector projects by prioritising usability and performance criteria over cost. This approach ensured that the selected team could deliver a system that met functional and usability requirements, highlighting how usability evaluation can effectively shape the Request for Proposals process. The case findings build on insights from the original study (anonymous citation for peer review), emphasising the value of integrating usability-focused practices into public procurement.

### 4.2 Case 2: The CFT for buying and developing further a CIMS Software

In contrast to case 1, case 2 examines the selection of a collaboration and information management system (CIMS) that had been developed. The selection process aimed at buying access to the selected system and the right to adjust the system to the users' needs and develop it further in collaboration with the selected software development company. The CIMS was to be used by all employees of all the ministries in a country in Northern Europe. The selection was done through defining an RTF, where criteria for selecting the right document management systems were defined.

#### 4.2.1 The collaboration and information management system

As stated in the formal RFP, that was a part of the description of the whole selection process, the goal of the procurement was to "*implement the System as an integrated information management system solution that is user friendly, widely in use and is actively being developed. The System should fulfil the needs and cover the procedures of Ministries as well as being economical to deploy and use at the Ministries. The System should enable the customer to fulfil legal requirements [which were referred to in detail]"*.

Regarding the digital work environment, it was stated that: "*With the implementation of the System, the Customer is aiming to improve the work environment at the Ministries, with the following emphasis:*

1. *The day-to-day work of employees and other users is made easier through modern solutions for access to information and collaboration, along with digitization of processes for all case processing.*
2. *Modern tools will provide excellent support for work processes and document management, including documentation of case processing and fulfilling legal requirements.*
3. *Modern tools will enable automation of work processes and regulated case processing*."

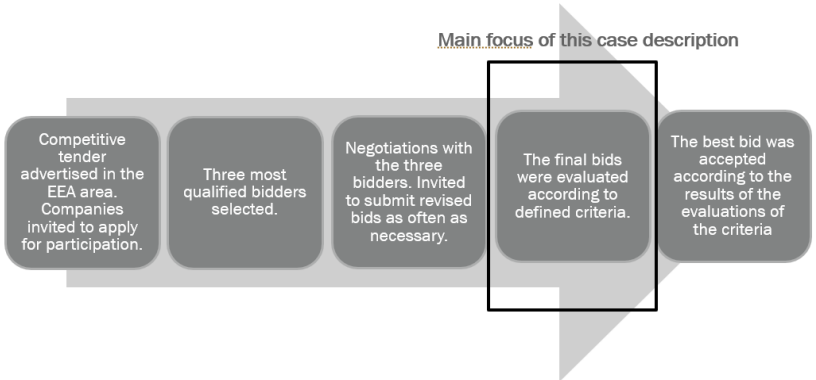
The estimated number of direct users was 850 persons at 10 ministries, and the number of external users in various committees and consuls was estimated to be 3.300 persons. The yearly number of cases was estimated to be 30.000, and the yearly number of new documents was estimated to be 350.000.

The requirements of the system were defined in several categories: General requirements, requirements for records management, requirements for workflow management, requirements for statistics, requirements for the user interface, requirements for collaboration, other requirements, and technical requirements. The requirements were grouped as Must (A) (total number = 85), Should (B) (total number= 3), and Could (C ) (total number none). One example of a general requirement was: "***Processes, projects and cases*** *- The System shall support the users in managing processes, executing projects and handling cases and required approvals*".

4.2.2 The Selection Process

The selection of the system was done through a five-step process, which is illustrated in Figure 1.

Figure 1: Illustration of the whole selection process




As can be seen in the figure, the tendering process was in five steps. First, the competitive tender was advertised in the EEA area. Six companies applied for

participation, four from Iceland and two from Denmark. In the second step, the bids were compared according to how well they fulfilled the baseline requirements. In that step, three companies were selected to be evaluated in more detail. In step three, negotiations were conducted with the three bidders, which were invited to submit revised bids as often as necessary.

When the final version of the bids had been handed in, the three bids were evaluated according to several criteria in step four. The selection criteria factors and weight of those are shown in Table 2.

**Table 2.** Selection factors and their weight.

| Selection criteria factor | | | Weight |
|---|---|---|---|
| **Cost** - estimation for implementation and operation for 10 years | | | 25% |
| **Quality** | | | 75% |
| Usability of the software: | | 20% | |
| Strength of the software: | 15% | | |
| Quality of the implementation team: | 10% | | |
| Quality of the project management process: | 20% | | |
| Special adaptions: | 10% | | |

We describe in more detail the process of evaluating the usability of the software in the selection process; see section 4.2.3.

In step five of the selection process, the best bid was accepted based on the evaluation results.

4.2.3 The Usability Evaluation Process

The system's usability was measured in user testing by observing how easy it was for users to solve particular tasks. The measured factors were effectiveness, by noting down the completion of the tasks, efficiency, by noting down the task completion time and satisfaction, by asking participants to fill in the SUS questionnaire.

The users were grouped into three groups: Managers, Document managers, and General users. Eighteen users took part in the usability evaluation, six from each user group. During the user testing, each participant used the three systems selected for 30 minutes each. Typically, the evaluation sessions lasted for 2 hours, with breaks in between.

Three conductors guided the participants in the usability evaluation, one of each user group. The conductors first introduced the evaluation process and then asked the participants to fill in a background questionnaire. Then the conductor handed the user's tasks to the participant to solve by using the particular system. During that session, the conductor noted the information on the effectiveness and efficiency. The conductor noted the question and how much help the user needed if the user wanted to ask questions. After each task-solving session, the conductor handed out the SUS satisfaction questionnaire for each system.

The tasks were based on the traditional handling of cases, traditional document types, and simple classification systems. Each task consisted of one example of the usage and

was defined to be realistic in the context of use for the users. Experts in documentation management were responsible for defining the tasks.

A pilot test was conducted as a rehearsal for the usability evaluation. The systems needed to be set up correctly at that point. One pilot user, who was not one of the expected users, was asked to take part in the pilot test. The pilot user went through the whole user testing process of all three systems, and the bidders were allowed to observe the pilot testing online. The contact persons observed how the user testing process was conducted during the pilot user testing. The contact persons were responsible for their systems working correctly during the pilot and actual testing. Through the pilot testing, they could ensure their system was set up and functioned correctly. The contact persons were not allowed to be present during the actual user testing. After the pilot testing, the bidders were allowed to make one A4 document as support material for their system for the usability evaluation.

#### 4.2.4 The Usability Evaluation Measures

The measured usability factors were effectiveness, efficiency, and satisfaction. The total usability rating was 100, where the weight of effectiveness and efficiency was 30% each and satisfaction was 40%.

The following formula gave points for effectiveness:

- 1 point: The task is regarded as solved satisfactorily (with the right data and finishing point).
- 0,5 point: The task was finished with the help of the conductor or with the wrong data.
- 0 points: The task was finished but estimated as unsatisfactory by the conductor of the user testing (because of unsatisfactory data or unsatisfactory completion).
- 0 points: The task was not finished within the time limit.
- 0 points: The user stopped solving the task before reaching the defined finishing point.

The effectiveness rating was calculated by summing all points given for each system for all users participating in the usability evaluation. The sum of points was divided by the number of tasks that users were asked to solve.

The calculation for the effectiveness rating was as follows:

$$\text{Rating for effectiveness for a particular system} = \frac{\text{The total number of points of the effectiveness from all users of solving all tasks in a particular system}}{\text{The total number of tasks that were handed out to all users using a particular system}} * \text{Effectiveness weight}$$

The efficiency was measured for each task, that the users solved in each system. The measurement started when the conductor handed out the tasks to the user and was stopped when the user indicated that he/she had finished the task. If a user did not finish a particular task for some reason, the task limit time was used as the task completion time, which was the maximum time for solving a task (10 minutes).

The calculation for the efficiency rating was as follows:

$$\text{Efficiency rating for a particular system} = \frac{\left(\text{The total sum of the time limits for all tasks for all users for a particular system} - \text{The total task solving time for all tasks for all users for a particular system}\right)}{\text{The total sum of the time limits for all tasks for all users for a particular system}} * \text{Weight of efficiency}$$

The satisfaction was measured by using the positive version of the SUS (System Usability Scale) questionnaire [24,25]. The positive version is recommended by the authors of the original SUS questionnaire [24]. The users were asked to answer the questionnaire after using each system. The rating for satisfaction was calculated according to the following formula:

$$\text{Satisfaction rating for a particular system} = \frac{\text{Total score from all users for a particular system}}{\text{The maximum score possible from all users}} * \text{Weight of satisfaction}$$

For calculating the total score for usability based on the results of the evaluations, the weighted ratings of effectiveness, efficiency and satisfaction were summed.

4.2.5 Usability Evaluation Results

The results from the usability evaluations of the three systems are shown in Table 3.

**Table 3.** The results of the usability evaluations

| Usability factor | System A | System B | System C |
|---|---|---|---|
| Effectiveness - *max 30 points* | 13,87 | 17,75 | 18,68 |
| Efficiency - *max 30 points* | 10,72 | 14,30 | 15,14 |
| Satisfaction - *max 40 points* | 13,72 | 22,78 | 22,72 |
| **Total - *max 100 points*** | **38,31** | **54,83** | **56,54** |

The results show that the usability of system B and C was quite similar, but system A got a much worse rating for usability. In the end system B was chosen, because other quality factors were stronger for that system than for system C.

4.2.6 Retrospective analysis of the evaluation process

After conducting the usability evaluation, the three conductors decided to gather their positive observations and ideas for possible improvements of the process. The conductors are referred to as C1, C2, and C3 when describing the results from the

retrospective meeting below. The three conductors met and discussed in a meeting room for one hour.

The main observations regarding the preparation of the usability evaluations were:

1. It would have been good to have a week, where all the systems were up and running before the actual usability evaluations were conducted. The first day was stressful since we felt the systems were not ready and we did not have enough practice on the evaluation process (C1, C3).
2. We were surprised at how long time it took to install the systems and make them run with the right data (C3).
3. It was good to have a skilled technical person, responsible for the installation of the systems for the usability evaluation (C1).
4. A more detailed definition of how a task is completed in each system, should have been made, to make it more clear for the conductors to rate the effectiveness in the usability evaluation (C2).
5. A more precise and coherent introduction to each system should have been made, like explaining what some of the buttons mean in each system (C2).
6. It was really good to do a pilot test before running the actual usability evaluations since that gave many insights both for the conductors and the bidding companies (C3).

The main observations regarding the conduction of the usability evaluations:

1. It would have been good to have an expert in usability evaluations observing the evaluation sessions since not all conductors were experts in usability evaluations (C1).
2. One of the tasks was too complex, being in three parts (C2).
3. A better testing of the tasks and what data was to be inserted in advance would have been better (C2, C3).

### 4.3 Synthesis of Case Study Findings

The two case studies demonstrate different approaches to incorporating usability evaluations into public procurement. Case 1 focused on evaluating development teams through a workshop-based selection process, assessing their collaboration, user-centred focus, and code quality. Case 2 involved structured usability testing of existing software, measuring effectiveness, efficiency, and user satisfaction. Each approach provided insights into usability integration in RFPs, highlighting differences in evaluation criteria and selection processes. These findings set the stage for further discussion on the implications of usability-focused procurement strategies.

## 5. Discussion

This study analysed two approaches to integrating work environment considerations into the Request for Proposals (RFP) process. In this section we will discuss and compare the two cases, we will discuss the limitations of the study and implications for advancing digital work environments.

### 5.1 Discussions of the two cases

The two case studies illustrate different approaches to integrating usability evaluation into public procurement: workshop-based team selection and structured usability testing of existing systems. While both methods effectively introduced usability criteria into the RFP process, they differ in scope, application, and suitability for different procurement contexts.

**Table 4.** Key Differences Between the Two Methods

| Aspect | Case 1:<br>Workshop-Based Selection | Case 2:<br>Structured Usability Testing |
|---|---|---|
| Purpose | Evaluating development teams for usability competence for developing new software | Assessing existing systems for usability and fitness for the digital work environment. |
| Assessment focus | Collaboration, user-centred design (UCD) focus, and code quality. | Effectiveness, efficiency, and satisfaction based on user testing. |
| Strength | Captures real-time collaboration, adaptability, and user focus | Provides quantifiable usability metrics with real users in realistic environment for decision-making. |
| Challenges | Resource-intensive, requires significant preparation and expert observation | Does not assess vendor adaptability or future usability improvements. Is also resource-intensive. |
| Best use case | When selecting a development team for a new system. | When evaluating existing software solutions in procurement |

The first case focused on selecting an agile development team for developing a financial support application, using a workshop-based methodology to evaluate team performance, collaboration, and adaptability. The second case examined procuring a collaboration and information management system (CIMS), relying on structured usability evaluation to assess the usability of three possible systems across diverse user groups.

In the first case, the workshop allowed evaluators to observe team interactions and their ability to align with user needs in real time. This approach was particularly beneficial for assessing the team's flexibility and capacity to adapt to the unique challenges presented by the project. However, the workshop methodology required significant preparation and participation, which could limit its scalability in large-scale procurement processes.

In contrast, the structured usability testing in the second case provided clear, measurable insights into how easy to use the preselected systems were for the various user groups. The evaluation captured valuable measurements on the systems' usability, effectiveness, efficiency, and satisfaction by involving representative users. The evaluations were conducted while running the actual systems and in as realistic environment as possible. However, this approach focused primarily on the usability of the systems. It did not evaluate the vendors' ability to adapt to the users' needs or collaborate effectively on making the systems more user centred, which might be critical in specific contexts.

The findings of this study align with prior research that emphasizes the advantages of evaluations as a part of public procurement. The workshop-based approach in the first case resonates with studies highlighting the importance of assessing collaboration and adaptability during vendor selection. For example, Tarkkanen and Harkke [10] stress that interactive evaluation methods, such as workshops, provide valuable insights into how teams work together and respond to user needs in real time. Similarly, Jokela et al. [8] argue that collaborative approaches allow procurement processes to focus on vendor flexibility and user-centred design, making them well-suited for custom development projects. However, this study also corroborates critiques regarding the resource-intensive nature of workshops, which limits their scalability in extensive or time-constrained procurement processes [26].

In the second case, structured usability evaluations support prior findings that this method generates precise and actionable data, particularly regarding effectiveness, efficiency, and user satisfaction [27] Such testing is especially effective in evaluating preselected systems where measurable metrics are needed to guide decision-making [9].

However, usability evaluations alone are not sufficient for procurement decisions. Specifically, structured evaluations do not address vendor adaptability or the ability to respond to broader organisational and work environment requirements, often critical in public sector contexts. This underscores the need for hybrid approaches that combine the strengths of collaborative and structured methods to achieve more comprehensive procurement outcomes.

### 5.2 Limitations

This study focuses on two case studies, aiming to provide insights that can be applied in similar contexts. While the detailed descriptions support transferability, the reliance on collaborative discussions rather than formal qualitative coding may limit methodological rigour. Furthermore, the study emphasises procurement processes and does not explore the long-term impact of these decisions on digital work environments. Future research could address these aspects to provide a more comprehensive understanding. Despite these limitations, the findings contribute perspectives on improving procurement practices to advance digital work environments.

### 5.3 Implications for Advancing Digital Work Environments

This study highlights the importance of tailoring procurement processes to effectively integrate considerations for digital work environments. The two cases demonstrate different approaches for integrating aspects of the users' digital work environment into the procurement processes, both before the systems are being developed and while selecting a running system, providing insights into how such procurement processes can meet organisational goals. In both these cases, the approach of the integration of work environment aspects was based on results from research studies, and the defined approach measuring the defined criteria was selected according to academic results and interpretations. We argue that such academic involvement leads to better procurement outcomes compared to traditional consulting approaches because

the academic approach builds on validated processes, and the benefits and drawbacks of those processes are known.

The findings suggest that integrating aspects of the users' digital work environment in the request for proposals (RFP) in public sector procurements is critical. In both cases described in the paper, the user needs and perspectives were given a considerable part of the selection criteria for choosing a development team for developing a usable system on the one hand and for choosing an already running usable system on the other hand. Defining the aspects of the digital work environment of the users in the procurement process will hopefully sharpen the focus on these critical aspects through the development and maintenance of the systems.

As time evolves, user needs often evolve, and digital work environments change, like can be seen clearly with the introduction of AI systems lately. The definitions of the digital work environment aspects in the RFP could act as a baseline for defining those aspects and allow organisations to adapt those aspects to changing requirements, ensuring long-term usability and relevance.

Future procurement strategies could benefit from combining elements of both approaches presented in this study. A hybrid model could provide a more comprehensive usability assessment while balancing resource constraints and procurement goals. Workshop-based evaluations can be valuable in early-stage procurement to assess vendors' user-centred design capabilities and collaborative problem-solving, while structured usability testing ensures that the chosen solution meets usability and efficiency requirements before implementation. By integrating qualitative and quantitative usability assessments at different stages of procurement, public sector organisations can improve system usability, create sustainable digital work environments, and enhance long-term user satisfaction.

These findings have important implications for public sector procurement. By integrating usability evaluation into tendering, this study offers a structured approach to improving digital work environments. Practitioners can adopt workshop-based or structured usability testing to enhance system selection. Future procurement should consider a hybrid approach to maximize usability and functionality.

## 6. Acknowledgements

This research was partly supported by Afa Försäkringar, Sweden [grant number 220244]. Additionally, we thank the participating researchers, practitioners, and organisations who contributed valuable insights and support throughout this work. Their collaboration has been instrumental in the successful completion of this research.

## References

1. Roto V, Law EC, Vermeeren AP, Hoonhout J. User experience white paper: Bringing clarity to the concept of user experience. 2011 [cited 2025 Jan 24]; Available from: https://research.tudelft.nl/en/publications/user-experience-white-paper-bringing-clarity-to-the-concept-of-us
2. Cajander Å, Larusdottir M, Hedström G. The effects of automation of a patient-

centric service in primary care on the work engagement and exhaustion of nurses. Qual User Exp. 2020;5(1):1–13.
3. Golay D, Salminen-Karlsson M, Cajander Å. Negative Emotions Induced by Work-Related Information Technology Use in Hospital Nursing. Comput Inform Nurs. 2021;
4. Tarafdar M, Tu Q, Ragu-Nathan TS, Ragu-Nathan BS. Crossing to the dark side: examining creators, outcomes, and inhibitors of technostress. Commun ACM. 2011 Sep;54(9):113–20.
5. Cajander Å. Usability – Who Cares? : The Introduction of User-Centred Systems Design in Organisations. 2010 [cited 2020 Feb 14]; Available from: http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-122387
6. Inal Y, Clemmensen T, Rajanen D, Iivari N, Rizvanoglu K, Sivaji A. Positive Developments but Challenges Still Ahead: A Survey Study on UX Professionals' Work Practices. J Usability Stud. 2020;15(4).
7. Jokela T. Determining usability requirements into a call-for-tenders: a case study on the development of a healthcare system. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries [Internet]. Reykjavik Iceland: ACM; 2010 [cited 2025 Jan 3]. p. 256–65. Available from: https://dl.acm.org/doi/10.1145/1868914.1868946
8. Jokela T, Laine J, Nieminen M. Usability in RFP's: The Current Practice and Outline for the Future. In: Kurosu M, editor. Human-Computer Interaction Applications and Services [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013 [cited 2025 Jan 24]. p. 101–6. (Hutchison D, Kanade T, Kittler J, Kleinberg JM, Mattern F, Mitchell JC, et al., editors. Lecture Notes in Computer Science; vol. 8005). Available from: http://link.springer.com/10.1007/978-3-642-39262-7_12
9. Billestrup J, Stage J, Larusdottir M. A case study of four IT companies developing usable public digital self-service solutions. In: The Ninth International Conference on Advances in Computer-Human Interactions [Internet]. 2016 [cited 2025 Jan 24]. Available from: https://www.academia.edu/download/81590864/download.pdf
10. Tarkkanen K, Harkke V. Evaluation for Evaluation: Usability Work during Tendering Process. In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems [Internet]. Seoul Republic of Korea: ACM; 2015 [cited 2025 Jan 24]. p. 2289–94. Available from: https://dl.acm.org/doi/10.1145/2702613.2732851
11. Solomou T, Mappouras S, Kyriacou E, Constantinou I, Antoniou Z, Canciu IC, et al. Bridging Language Barriers in Healthcare: A Patient-Centric Mobile App for Multilingual Health Record Access and Sharing. Front Digit Health [Internet]. 2025 Jan 31 [cited 2025 Feb 6];7. Available from: https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2025.1542485/abstract
12. Cajander Å, Sandblad B, Stadin M, Raviola E. Artificial intelligence, robotisation and the work environment Literature review. Swedish Agency for Work Environment Expertise; Report No.: 978-91-987761-8–8.
13. Roto V, Palanque P, Karvonen H. Engaging automation at work: a literature review. Human Work Interaction Design. Designing Engaging Automation: 5th

IFIP WG 13.6 Working Conference, HWID 2018, Espoo, Finland, August 20-21, 2018, Revised Selected Papers 5. Springer; p. 158–72.
14. Karltun A, Karltun J, Berglund M, Eklund J. HTO – A complementary ergonomics approach. Appl Ergon. 2017 Mar;59:182–90.
15. International organisation for standardisation: ISO 9241-210:2019. Ergonomics of human-system interaction - Part 210: Human-centred design process for interactive systems, 2019.
16. Nielsen J, Molich R. Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI conference on Human factors in computing systems Empowering people - CHI '90 [Internet]. Seattle, Washington, United States: ACM Press; 1990 [cited 2025 Jan 24]. p. 249–56. Available from: http://portal.acm.org/citation.cfm?doid=97243.97281
17. Boren T, Ramey J. Thinking aloud: Reconciling theory and practice. IEEE Trans Prof Commun. 2000;43(3):261–78.
18. Ha S, Ho SH, Bae YH, Lee M, Kim JH, Kim JH, et al. Digital Health Equity and Tailored Health Care Service for People With Disability: User-Centered Design and Usability Study. J Med Internet Res. 2023;25:e50029.
19. Larusdottir MK. Usability Evaluation in Software Development Practice. In: Campos P, Graham N, Jorge J, Nunes N, Palanque P, Winckler M, editors. Human-Computer Interaction – INTERACT 2011. Berlin, Heidelberg: Springer; 2011. p. 430–3. (Lecture Notes in Computer Science).
20. Larusdottir MK. User Centred Evaluation in Experimental and Practical Settings [PhD Thesis]. KTH, Media technology and interaction design; 2012. (Trita-CSC-A).
21. Larusdottir M, Cajander Å, Gulliksen J. Informal feedback rather than performance measurements: User-centred evaluation in Scrum projects. Behav Inf Technol. 2014;33(11):1118–35.
22. Larusdottir MK, Gulliksen J, Hallberg N. RAMES–Framework supporting user centred evaluation in research and practice. Behav Inf Technol. 2019;38(2):132–49.
23. Vindrola-Padros C, Johnson GA. Rapid Techniques in Qualitative Research: A Critical Review of the Literature. Qual Health Res. 2020 Aug;30(10):1596–604.
24. Sauro J, Lewis JR. When designing usability questionnaires, does it hurt to be positive? In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems [Internet]. Vancouver BC Canada: ACM; 2011 [cited 2025 Feb 5]. p. 2215–24. Available from: https://dl.acm.org/doi/10.1145/1978942.1979266
25. Brooke J. SUS-A quick and dirty usability scale. Usability Eval Ind. 1996;189(194):4–7.
26. Larusdottir MK, Kyas M. Selecting the Best Agile Team for Developing a Web Service. In: Human Computer Interaction and Emerging Technologies: Adjunct Proceedings from the INTERACT 2019 Workshops. Cardiff University Press Cardiff, UK; 2020. p. 289–302.
27. Jaspers MW, Steen T, Van Den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inf. 2004;73(11–12):781–95.