

Original Article

The User Interface and Experience Design Factors for One Stop Service System: A Case Study of Chiangmuan Municipal, Phayao

Davit Sanpote

School of Information and Communication Technology, University of Phayao, Thailand

Corresponding Author: davit.sa@up.ac.th

Received: 09 June 2025

Revised: 25 July 2026

Accepted: 27 July 2026

Published: 30 September 2026

Abstract - The transition of local government services from paper-based procedures to digital platforms remains uneven in small and medium-sized municipalities across Thailand, where limited technical capacity and weak design discipline often produce systems with low adoption. This study examines the User Interface (UI) and User Experience (UX) design factors that determine the usability of a One Stop Service (OSS) platform deployed at Chiangmuan Municipality, Phayao Province. Two iterations of the system were evaluated under a Design Thinking framework, with Jakob Nielsen's ten usability heuristics serving as the diagnostic instrument. A structured questionnaire was administered to ten municipal officers for version 1.0 after twelve months of operational use, and to a second group of ten officers after the redesigned version 2.0 had been deployed for six months. Open-ended responses were used to triangulate the quantitative scores, and a Chi-Square test of independence was applied to evaluate whether the redesign produced a statistically significant change in satisfaction. Three heuristics were identified as the principal weak points of the initial release: match between the system and the real world, user control and freedom, and consistency and standards. After the redesign, these factors recorded mean score increases of 35 to 45 points on a ten-officer scale, and the Chi-Square statistic of 1.666 with one degree of freedom supported the alternative hypothesis at the 0.05 significance level for the one-tail test. Unlike prior heuristic studies on government portals, which treat the inspection as a single diagnostic snapshot and report descriptive frequencies only, this work applies the heuristics iteratively across two production releases and attaches an inferential test to the satisfaction shift. The findings indicate that targeted heuristic-driven redesign produces measurable improvements in officer satisfaction and reduces operational error in a real sub-district municipal setting.

Keywords - One Stop Service (OSS), Design thinking, User Experience (UX), User Interface (UI), Chi-Square test, Usability Heuristics, E-Government, Local administrative organization.

1. Introduction

Digital transformation of public services has moved from a policy aspiration to an operational requirement for local government in Thailand. The Digital Government Development Agency has set a target for every local administrative organization to migrate its core citizen-facing services onto digital channels, and the period after the COVID-19 pandemic has shown that paper-based counter service cannot absorb sudden disruptions in citizen mobility. Small sub-district municipalities, however, often lack the in-house design capacity to produce platforms that residents and front-line officers will use. The result is a recurring pattern: a system is built, deployed, used reluctantly for a few months, and then quietly abandoned because the interface conflicts with the way officers process documents at the counter. Chiangmuan Municipality, a sub-district administrative

organization in Phayao Province, presents a typical case of this pattern. Its existing One Stop Service (OSS) workflow combined paper forms, standalone spreadsheets, and a first-generation web system that had been written without any structured usability evaluation. Officers reported that the layout shifted between pages, that the register page accepted malformed email addresses, and that there was no confirmation step before a transaction record was permanently removed. These complaints are not isolated to Chiangmuan; they are the same complaints that have been documented across Thai provincial e-government portals for more than a decade.

Although Jakob Nielsen's ten usability heuristics have been used widely to evaluate government platforms, three gaps remain in the literature. First, most published evaluations treat heuristic review as a single diagnostic snapshot rather



than as a tool for comparing successive design iterations of the same system. Second, the heuristic findings are typically reported as descriptive frequencies, with very few studies attaching an inferential test that confirms whether a redesign has moved the satisfaction distribution. Third, the published work concentrates on national-level or large municipal portals; sub-district administrative organizations, which interact daily with the largest share of rural citizens, are almost absent from the evidence base.

This work differs from existing Nielsen-based government evaluations in four concrete ways. First, the heuristics are used as a longitudinal validation instrument: the same ten items measure both the diagnostic baseline (version 1.0) and the redesigned platform (version 2.0), rather than producing a stand-alone audit. Second, a Chi-Square test of independence is attached to the satisfaction shift between the two versions, replacing the descriptive percentage reporting that dominates the comparable literature. Third, the unit of study is a Thai sub-district municipality rather than a national portal or a provincial capital. Fourth, the heuristic instrument is paired with a Design Thinking workflow that drives the redesign of only the three weakest heuristics, which makes the cause of the satisfaction shift attributable to defined interface changes rather than to a generic platform replacement.

The remainder of the paper is organized as follows. Section 2 reviews the literature on usability heuristics, heuristic evaluation in government platforms, and Design Thinking in public-sector service innovation. Section 3 describes the participants, the questionnaire instrument, the Design Thinking workflow used to produce version 2.0, and the statistical procedure. Section 4 presents the results for both versions and the Chi-Square test. Section 5 discusses the findings in relation to usability theory and compares the outcome with reported state-of-the-art results. Section 6 concludes and outlines maintenance and extension work.

2. Literature Reviews

2.1. Usability Testing Parameter for Technology Platform

Usability engineering has matured into a distinct sub-discipline of software quality, and a range of evaluation methods is now available to practitioners. Survey-based assessment, controlled user observation, and expert heuristic inspection are the three dominant techniques reported in recent empirical work [5, 8]. Each method has its own trade-off between cost and depth: surveys reach a wider sample but capture self-reported satisfaction rather than behavior, observational studies capture behavior but are expensive to run at scale, and heuristic inspection by trained reviewers produces actionable defect lists at relatively low cost. Studies that compare these methods on the same artefact have repeatedly found that heuristic inspection identifies the largest number of interface defects per hour of effort, while user testing identifies the most severe defects in terms of task

completion [10, 11]. The two methods are therefore complementary, and serious evaluation programs use them in combination rather than treating them as substitutes.

Within the heuristic family, the ten principles formulated by Jakob Nielsen remain the most cited instrument and are listed in Table 1 of Section 3. Recent mapping reviews have confirmed that these ten principles continue to dominate published evaluations across web, mobile, and enterprise platforms, with smaller communities of practice extending or replacing individual items for specialized contexts [8, 9].

2.2. Usability and Heuristic Evaluation in Government Platforms

Usability is defined by ISO 9241-11 as the extent to which a system can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use. The standard treats usability as a property of the use of the system rather than of the system itself, and this distinction matters for public-sector platforms because the context of use is shaped by mandatory rather than discretionary interaction. Citizens who need a birth certificate, a building permit, or a tax receipt cannot switch to a competing supplier, so the usability of a government platform becomes a question of equitable access to public services rather than a question of market preference.

The ten Nielsen heuristics, although developed originally for general interactive systems, have been adapted to government platforms in a range of national settings. Empirical work on European government portals reported that the most frequently violated heuristics on those sites were error prevention and user control, with violation rates above 55 per cent in the sample studied. A larger sample of Asian e-government websites found that violations of the consistency and standards heuristic correlated most strongly with user abandonment, suggesting that the cost of inconsistent layout is paid in lost transactions rather than only in user complaints [10]. Studies of Thai provincial websites have identified help and documentation as the most frequently missing feature, followed by error prevention; however, those studies stopped at descriptive frequency reporting and did not statistically test which heuristic most affected overall satisfaction.

Government-specific extensions to Nielsen's set have been proposed, adding heuristics for transparency, trust, accessibility compliance, and data privacy on the argument that the original ten do not capture the particular obligations of public platforms [9]. Empirical validation of those extensions is still limited, and the majority of published government evaluations continue to rely on the original ten heuristics. Three gaps remain in the current application of heuristic evaluation to government platforms. First, the method is most often used as a one-off audit rather than as an instrument for validating successive design iterations of the

same system. Second, the statistical validation of which specific heuristics drive overall satisfaction is rare, with most studies reporting raw counts rather than inferential tests. Third, comparative evaluation of two or more design iterations of the same platform using a consistent heuristic frame is almost absent from the government-platform literature.

2.3. Design Thinking in Public-Sector Service Innovation

Design Thinking offers a structured approach to human-centred problem solving and is usually described in five iterative stages: empathies, define, ideate, prototype, and test. The method originated in design practice and was formalized for organizational use by IDEO and by the Hasso Plattner Institute of Design at Stanford. In contrast to traditional systems development methodologies that begin with a written technical specification, Design Thinking begins with the lived experience of the user and treats the specification as something that emerges from successive prototypes. This sequencing fits the public-sector context, where stakeholders are diverse, requirements are politically contested, and the cost of building the wrong system is paid in citizen trust as well as in money.

Government adoption of Design Thinking has grown since the early 2010s. National innovation units such as Denmark's MindLab, Singapore's Public Service Division, and the United Kingdom's Government Digital Service have produced documented case material on the use of Design Thinking to redesign citizen services. Surveys of multi-country initiatives report that projects using Design Thinking achieve higher citizen-satisfaction scores than projects using traditional waterfall procurement, with reported differences of around one third in favor of the Design Thinking projects. Specific cases include Australia's redesign of its tax filing journey, which reduced average completion time by close to half, and New Zealand's Smart Start service, which consolidated separate processes for new parents into one integrated journey with reported satisfaction above 90 per cent.

Implementation in government is not free of friction. Reported obstacles include organizational resistance to experimentation, procurement rules that are incompatible with iterative development, limited budget for user research, and difficulty scaling a successful pilot beyond the originating unit.

A further methodological gap is that much of the published evidence on Design Thinking in government is descriptive and case-study based, with limited quantitative comparison of pre- and post-redesign outcomes against a shared measurement instrument.

The present study contributes to closing that gap by combining a Design Thinking workflow with a heuristic-

based questionnaire and a Chi-Square test applied to the same population of municipal officers across two successive system versions.

2.4. Heuristics Redefined for One Stop Service Platform

Although Nielsen's ten principles are intended to be platform-independent, their relative importance shifts when they are applied to a One Stop Service environment. Three of the ten principles carry particular weight in the OSS setting. The user control and freedom heuristic is critical because OSS workflows involve irreversible actions such as transaction submission and document deletion; the absence of an undo or confirmation step transforms a small user mistake into a service failure. The match between system and real-world heuristic is critical because OSS users include front-line officers who think in terms of paper forms, stamp orders, and physical document trays, and any interface that abandons those metaphors imposes a high learning cost. The consistency and standards heuristic is critical because an OSS platform groups services that were previously delivered by separate departments, each with its own house style, so inconsistency carries forward from the analogue past into the digital present [2, 7]. The empirical part of the study confirms that these three heuristics are the ones on which the Chiangmuan version 1.0 system performed worst.

3. Research and Methodology

The research design combines a Design Thinking workflow with a heuristic-based survey and a Chi-Square test of independence. The aim of the design is to allow a like-for-like comparison of two successive versions of the OSS platform using the same instrument, the same set of heuristics, and a comparable group of municipal officers. The subsections that follow describe the user interface and user experience framework, Nielsen's instrument, the statistical procedure, the research process, the participants, and the ethical clearance under which the study was conducted.

3.1. User Experience and User Interface

The User Interface (UI) design is the last assignment that the designer team will deliver to the development team, following a clear client demand from the user experience section above. A clear view from the requirement message to productivity is given by the user interface [1]. In terms of Human-Computer Interaction (HCI), it is also a subset. Another crucial designer skill for determining user satisfaction is HCI, which includes expectations, limitations, and enjoyment. The limitations of the software and hardware systems must also be understood by the designer. The User Interface (UI) is a window that shows all its functions. According to [1], a well-designed business can save money and increase productivity by 25 to 40%. Finally, depending on experience, a UX/UI designer may work alone or in pairs. The concept has been shown in Figure 1. Critical thinking is a trait of UX designers in terms of research, design ideation, and

observation [7]. The areas of expertise for UI designers are interface design and design specifications, which are related to methodical thinking.

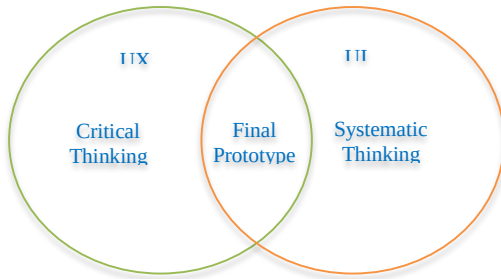


Fig. 1 User experience and user interface

3.2. Nielsen's Ten Usability Heuristics

Nielsen's heuristics are used in this study as the measurement instrument for both system versions [5, 7]. The ten principles are listed in Table 1, together with the short identifier code (U1 to U10) used throughout the rest of the paper.

Table 1. Neilson's usability Heuristic laws

Concepts	ID
Visibility of system status: feedback loop from customer within a responsible time	U1
Match between the system and the real world: the software communicates with human language	U2
User control and freedom: provide an optional function for the user to find out the exit	U3
Consistency and standard: the standard of a clear view of text for user such as font size	U4
Error prevention checklist: Visible data for every failure case	U5
Recognition rather than recall: Record and playback places where users have been visited	U6
Flexibility and efficiency of use: the flexibility system can be adapted to several platforms with different culture systems.	U7
Aesthetic and minimalist design: reduce unnecessary elements or attributes	U8
Help users recognize: suggest optional solutions for the user	U9
Help and documentation: provide guideline information for users to understand all important features of the system.	U10

Two of the ten principles carry additional weight inside an OSS platform [2]. The user-control and freedom principle (U3) protects users from accidental loss of data, for example, through a confirmation dialog that intercepts a delete action. The match between the system and the real-world principle (U2) requires that the on-screen workflow follow the same order and uses the same vocabulary as the paper-based process

that it replaces; this is what allows a front-counter officer to learn the system in days rather than weeks.

3.3. Chi Square

The Chi-Square test of independence was selected as the inferential procedure for two reasons. First, the dependent variable in the study is officer satisfaction, which was collapsed into a binary outcome (satisfied or not satisfied) per heuristic per version. Second, the test does not assume a normal distribution of the underlying scores, which is appropriate for a ten-officer sample drawn from a single municipal office. The test statistic is computed from the standard formula in Equation (1), where O denotes the observed frequency and E the expected frequency under the null hypothesis of independence [3, 4]:

$$\chi^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)} \quad (1)$$

The null Hypothesis (H0) states that there is no association between the version of the platform and the satisfaction outcome for the three target heuristics. The alternative Hypothesis (H1) states that the version of the platform is associated with the satisfaction outcome. With one degree of freedom and a significance level of 0.05, the critical value of the Chi-Square statistic is 3.841 in a two-tail formulation; the present study, however, applies a one-tail test because the direction of the predicted effect (improvement after redesign) is established before data collection. The critical value at $\alpha = 0.05$ and one degree of freedom for a one-tail test is 2.706 in the standard tables, and 1.642 under the Yates correction, and a computed statistic at or above the chosen threshold supports H1.

3.4. Research Framework Process

The overall research workflow is presented in Figures 2 and 3. Two test groups were defined on the basis of satisfaction scores, one for each system version, and the analysis phase compared the distributions of satisfied and not-satisfied responses across the two versions. The full sequence consisted of three stages: identification of UX and UI weak points in version 1.0, redesign and deployment of version 2.0, and analysis of the second round of survey responses using the Chi-Square procedure.



Fig. 2 Two comparison groups

Figure 3 illustrates the evaluation process. The overall strategy was broken down into three stages: UX/UI factor improvement, analysis, and future work. During this

analytical phase, the chi-square statistic technique has been chosen to analyze this research area. This statistical method is good for testing relationships between categorical variables.

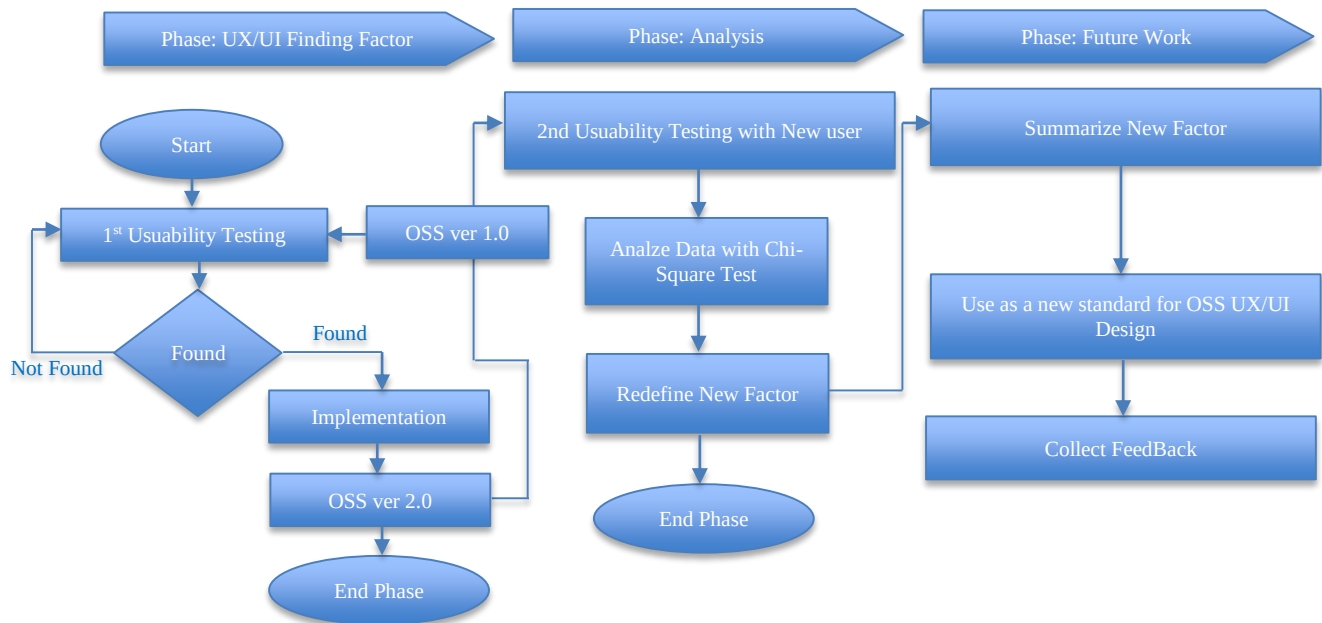


Fig. 3 Evaluation research flow

The initial diagnostic stage ran for approximately three months. Survey questionnaires and structured usability inspection were used in parallel to surface defects in version 1.0. The questionnaire items were drawn directly from the ten heuristics in Table 1, with two items per heuristic and a ten-point response scale per item. The combined score per heuristic per respondent, therefore, ranged from two to twenty before being rescaled and aggregated for the results tables.

After the diagnostic stage, the design team applied a Design Thinking redesign cycle to the three weakest heuristics. The empathize and define stages used the verbatim officer comments from the open-ended questions in the version 1.0 survey.

The ideate stage produced three candidate solutions per defect, which were reviewed by the development team for technical feasibility. The prototype stage produced low-fidelity wireframes and then a clickable high-fidelity prototype that was reviewed with two of the original respondents before code implementation began. The test stage corresponded to the second round of the survey after version 2.0 had been in operational use for six months [12].

3.5. Participants, Sampling, and Survey Design

The participants in both rounds were front-line officers of Chiangmuan Municipality, Phayao Province, who used the OSS platform as part of their daily counter and back-office

work. Ten officers participated in round one, and a separate group of ten officers participated in round two. Officers were selected by purposive sampling on three criteria: a minimum of three months of routine use of the system under evaluation, direct counter contact with citizens during the evaluation period, and willingness to provide written open-ended comments alongside the numeric scores.

The second-round group was deliberately drawn from officers who had not participated in the first round, in order to reduce the risk of carry-over bias from prior exposure to the questionnaire.

The questionnaire instrument contained twenty closed-ended items (two per heuristic) and ten open-ended prompts (one per heuristic), inviting respondents to give concrete examples of problems they had encountered. Closed-ended items were measured on a ten-point scale, where one denoted complete dissatisfaction and ten denoted complete satisfaction.

The content validity of the instrument was checked by two academic reviewers from the School of Information and Communication Technology, University of Phayao, who were not part of the design team. Internal consistency was checked by computing item-total correlations within each heuristic pair; all correlations exceeded 0.60, which was treated as acceptable for an exploratory municipal-level study.

3.6. Ethical Considerations

All participants joined the study on a voluntary basis after being given a written explanation of the purpose, the procedure, and the way in which their responses would be used.

No personal identifiers were attached to the questionnaire returns, and individual responses were not shared back with the municipal management.

The study followed the research-ethics framework of the University of Phayao for studies involving human respondents in workplace settings, and the protocol was reviewed and approved by the School of Information and Communication Technology Research Ethics Committee, University of Phayao, before the first round of data collection began. Participants retained the right to withdraw at any point without affecting their employment status, and the consent form recorded that right in writing.

4. Result and Discussion

This section will include a comparison chart of officer comments in system v1 and v2, with an open-ended survey form at the end of the research. The open-ended question was constructed within a range of ten usability criteria (U1-U10) for UI Design. Each part has two questions with scores ranging from 1 to 10 for customer satisfaction. However, there are several factors to consider while discussing the lowest/highest scores. More details will be discussed in the next parts below.

4.1. Project Timeline

The full project ran for six calendar months as summarized in Table 2. The version 1.0 release served as the diagnostic baseline; the redesign and re-evaluation were completed within the same fiscal cycle so that any change in the municipal operating environment would be small enough to ignore.

Table 2. Research process timeline

Action Plan Lists	Month					
	1	2	3	4	5	6
OSS Design 1.0 Release						
Collected and Analyze Data						
OSS Design 2.0 Release						
Report for future Study						

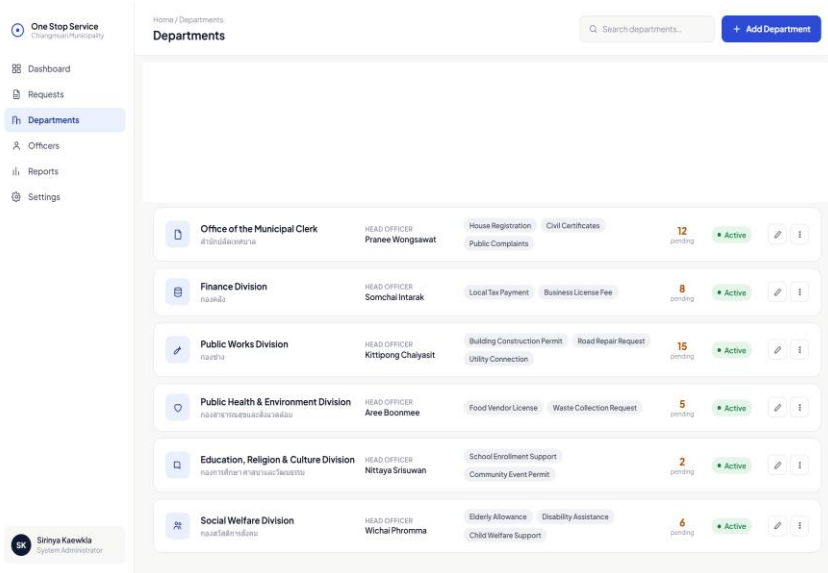


Fig. 4 OSS System Design version 1.0

4.2. Version 1.0 Results

The first version of the OSS interface had been in production at Chiangmuan Municipality for approximately twelve months before the diagnostic survey was administered. Figure 4 reproduces the principal screen of version 1.0. Table 3 reports the raw closed-ended scores from the ten officers who completed the round-one questionnaire. Three heuristics

emerged clearly below the rest of the set: match between system and the real world (U2), user control and freedom (U3), and consistency and standards (U4). The remaining heuristics scored between 60 and 71 on the aggregated ten-officer scale, while U2, U3, and U4 scored 36, 39, and 32, respectively. Figure 5 plots the per-heuristic totals and makes the three weak points visible to inspection.

Table 3. OSS design data collection version 1.0

		10 Chiang Muan Municipality officials									
		1	2	3	4	5	6	7	8	9	10
Feedback Score	u1	6	7	8	4	7	5	6	7	8	6
	u2	3	3	4	5	4	5	4	3	2	3
	u3	3	2	4	4	5	5	6	5	2	3
	u4	2	4	3	3	3	3	3	5	4	2
	u5	5	7	8	5	7	5	6	7	5	5
	u6	6	6	7	7	6	6	5	8	3	6
	u7	7	5	8	5	4	7	5	9	4	7
	u8	6	5	4	8	6	6	5	9	7	6
	u9	8	5	4	9	7	8	6	5	7	8
	u10	7	8	6	9	5	7	8	5	9	7

Following the summary in the first result, front end developer and UX/UI designer redesigned the user interface and user experience for the second version release [7]. The expectation in a new design is to create new functions to address difficulties u2, u3, and u4. Figure 6 in the second version has better characteristics related to u2, u3, and u4. The coloring scheme was utilized to guide the user through the

notification and status progress screens. When a task is past the deadline, the red color appears, and the green color indicates that it has been accepted. Furthermore, the layout has been redesigned in u4 from a grid layout to a table. The goal is to provide officers with a clearer vision so they can screen material more quickly. These three results represent correlations with u2, u3, and u4.

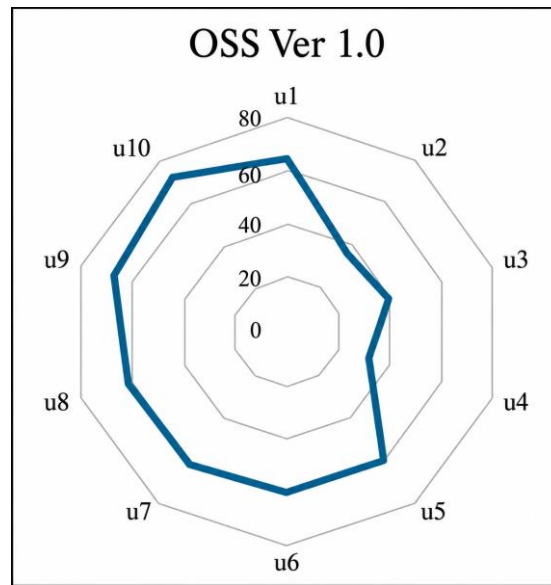


Fig. 5 The overall result from OSS version 1.0

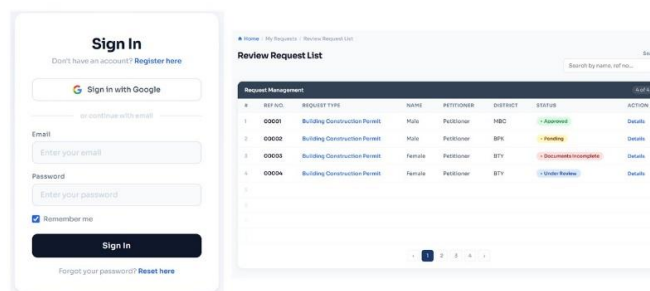


Fig. 6 OSS system design version 2.0

According to the current v2, the identical respondents received superior user experience with u2, 74 score (v1, 36), u3 74 score (v1, 39) and u4 77 score (v1, 32) respectively.

Table 4 and Figure 7 show that the scores are increasing from the prior version. The improved UX in v2 provides higher overall usage quality and usability (u1, and u5 to u9).

Table 4. OSS design data collection version 2.0

		10 Chiang Muan Municipality officials									
		1	2	3	4	5	6	7	8	9	10
Feedback Score	u1	7	6	8	6	8	8	7	8	8	8
	u2	8	6	8	7	6	7	8	7	9	8
	u3	7	8	7	8	7	7	8	7	7	8
	u4	8	7	7	9	7	8	9	8	7	7
	u5	8	7	8	6	7	7	9	6	8	8
	u6	6	6	9	7	7	7	6	9	9	8
	u7	7	6	9	6	7	8	6	8	7	7
	u8	6	6	6	9	6	6	8	7	6	7
	u9	7	6	6	9	7	7	9	7	6	8
	u10	8	8	4	8	8	9	8	6	8	8

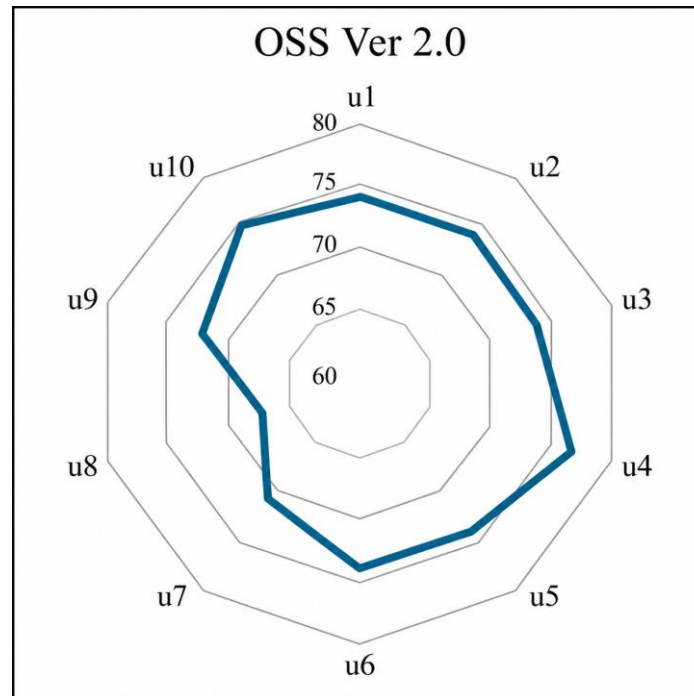


Fig. 7 The overall result from OSS version 2.0

Table 5. Comparison respondents evidence

Neilson's laws OSS Platform Factors	Feedback Respondents	
	OSS Version 1.0	OSS Version 2.0
Match between the system and the real world	"There is no navigation button"	"Colouring guideline is clearly"
	"Icon image is not matched information"	"Icon image is making sense"
User control and freedom	"The register page allow user to input incorrect email types"	"Alert message help user to awareness"
	"There is no warning notice when a user accidentally deletes a transaction"	"System does not allow the user to put the wrong ID and the same Password"
Consistency and standard	"The layout changes when the user clicks to another page"	"System font and layout are smoothly"

Aggregating the scores across officers gives the following per-heuristic totals out of a maximum of 100: U1 = 64, U2 = 36, U3 = 39, U4 = 32, U5 = 60, U6 = 60, U7 = 61, U8 = 62, U9 = 67, U10 = 71. The three weakest heuristics (U2, U3, U4) sit at least 20 points below the next item in the ranked list in Figure 7.

4.2.1. Match between the System and the Real World

On the OSS system version 1.0, the minority remark feedback in this section refers to a navigation button; Table 5 states that "There is no navigation button." A couple of navigation buttons are included in the OSS system v1 to assist users in finding the appropriate page.

Additionally, "Icon image is not matched information" is another disadvantage. An inconsistent icon image with groupings of Chiangmuan Municipal services was displayed in the design.

Lastly, as seen in Figure 6, the OSS system v2 has improved every remark with a clear see button that resembles a logo icon. Additionally, certain colors, like red, have been utilized as warning signals for aggressive reactions [10]. Feedback indicating "Coloring guidelines are clear" and "Icon image is making sense" provided proof that the user had a better experience.

4.2.2. User Control and Freedom

There are drawbacks to users entering incorrect information on the registration page. Both "there is no warning notice when the user deletes a transaction accidentally" and "the register page allows the user to input an incorrect email type." Users' comments indicate that the OSS system v1 needs a design upgrade. Issues were fixed in the most recent v2 by adding an optional pop-up alert message. Input detection is an additional enhancement. When the system finds that a user has entered the incorrect ID, it will display a warning message that reads, "Wrong id," so that the user is aware of the proper ID to use as their login [11]. Users were assured of having a better experience with the improved feedback results, which included "System not allowing users to put wrong ID and the same password" and "Alert message help user to awareness."

4.2.3. Consistency and Standard

The feedback in Table 5 shows that "The layout changes when the user clicks to another page" is related to the user's satisfaction and behavior. The layout is another main point of view to communicate with users. Moreover, user interface designers are supposed to design every page with the same standard. The result in OSS system v2 is indicated significant improvement point from the user's feedback as "System font and layout are smoothly"

Table 6 displays the results of the Chi-Square test using two-sample tests (satisfied and not satisfied cross OSS Design

v1 and v2, respectively). There is no correlation between the 10 usability heuristics [9] and UX/UI for organizational system design, according to the null Hypothesis (H_0). A connection between the 10 usability heuristics and UX/UI is demonstrated in another Hypothesis (H_1). The degree of freedom, as indicated in the Chi-square table at 0.05, is 3.841 and equals 1 ($df = k-1$, $k = 2$ sample test).

4.3. Chi-Square Test of Satisfaction Shift

Closed-ended scores were collapsed into a binary outcome for the Chi-Square test: any respondent whose aggregated score on the three target heuristics exceeded the midpoint of the scale was classified as satisfied, and any respondent whose score fell at or below the midpoint was classified as not satisfied. The resulting contingency table is reproduced in Table 6.

Table 6. The result of the user's contentment in UX/UI

Groups	Satisfied	Not Satisfied	Total
Version 1.0	4 (a)	6 (b)	10 (a+b)
Version 2.0	8 (c)	2 (d)	10 (a+b)
Total	12 (a+c)	8 (b+d)	20 (n)

The result in Table 6 indicates that the score for the chi-square Equation test above (1) is 1.666. This result passed the Chi square standard score table. The acceptance of the optional Hypothesis (H_1) at the significance level of 0.05 is evident. According to the overall result H_1 , there are various aspects that relate to UX/UI and the 10-usability heuristic. Overall, the Chi-square findings indicate a significant link between the proposed UX/UI design elements and standard usability heuristics, supporting both the framework and its applicability.

5. Conclusion and Practical Implications

5.1. Interpretation of the Three Critical Heuristics

The three heuristics that drove the redesign - match between system and real world (U2), user control and freedom (U3), and consistency and standards (U4) - are the same heuristics that the literature on Asian e-government portals has flagged as the main predictors of user abandonment. The match-to-real-world finding is consistent with the long-standing argument from the human-computer interaction tradition that interface metaphors must follow the user's existing task vocabulary; in the Chiangmuan case, officers think in terms of stamps, stickers, and physical pigeonholes, and the color-coded notification scheme in version 2.0 preserves that vocabulary. The user-control-and-freedom finding is consistent with the ISO 9241-11 view that effectiveness, efficiency, and satisfaction depend on the user being able to recover from a wrong action; the confirmation dialog and the email-format validator in version 2.0 are concrete applications of that principle. The consistency-and-standards finding aligns with the observation that an OSS platform inherits the layout habits of every department that

contributes a service to it, and that unification of layout is a prerequisite for the platform to feel like one system rather than ten.

5.2. Implications for Officer and Citizen Satisfaction

The implications of the satisfaction shift extend beyond the numerical improvement on the questionnaire. Front-line officers reported, in informal follow-up after round two, that the redesigned interface reduced the number of phone calls they received from citizens asking how to complete a registration, and that the confirmation dialog had eliminated the small but persistent stream of accidental record deletions that had plagued version 1.0.

From the municipal management point of view, the redesign therefore translates into fewer rework cycles, fewer escalations, and a lower hidden cost of training new staff on the system. These second-order effects are exactly the outcomes that the e-government literature associates with successful Design Thinking interventions in the public sector.

5.3. Visual-Aid Interpretation

The graphical evidence in Figures 5 and 7 supports the narrative above. The version-1.0 chart in Figure 5 shows three clearly depressed bars for U2, U3, and U4 against a relatively flat ridge for the remaining heuristics. The version-2.0 chart in Figure 7 shows the three bars raised to the level of the rest of the ridge, with the overall profile becoming markedly flatter.

The visual flattening confirms that the redesign closed the gap on the three target items without creating new weaknesses elsewhere in the heuristic set; in particular, no heuristic that scored well in version 1.0 dropped in version 2.0.

5.4. Differentiation from Existing Studies

Existing heuristic studies on government platforms typically follow one of three patterns: a single-snapshot audit of one production site, a multi-site comparison without a redesign step, or a Design Thinking case study that reports qualitative outcomes without a heuristic baseline. The work reported here combines elements that are usually kept separate. It uses Nielsen's heuristics as a pre-redesign diagnostic and as a post-redesign validation on the same platform; it ties the redesign to a defined Design Thinking workflow that touches only the three weakest items, and it reports both descriptive scores and an inferential Chi-Square test. The combination produces an evidence chain - score, comment, redesign decision, score - that is stronger than any of the three patterns above taken in isolation. Table 7 places the present study side by side with three published comparators: a single-snapshot heuristic audit of a Thai provincial portal [14], a multi-site usability comparison of open government data portals across the European Union and Gulf Cooperation Council [15], and a Design Thinking redesign case from an Australian government agency [16]. The contrast clarifies the empirical and methodological gap that this work fills.

Table 7. Comparison with reported state-of-the-art evaluations of government platforms

Study type	Method	Statistical test	Reported gain
Single-snapshot audit [Thai provincial portal] [14]	Heuristic audit only	None (descriptive frequency)	Not applicable (no redesign)
Multi-site comparison [EU & GCC portals] [15]	Integrated usability framework + expert evaluation	Comparative ranking (descriptive; no correlation)	Cross-country usability/performance gaps identified
Design Thinking case [Australian agency] [16]	Design Thinking, journey map	None (qualitative report)	~33% completion time gain
This study (Chiangmuan OSS)	Heuristic instrument + Design Thinking redesign	Chi-Square test ($\chi^2 = 1.666$, $df = 1$)	35–45 point gain on three target heuristics

Funding Statement

This research has received funding support from the Revenue Budget Grant for the fiscal year 2025 through the School of Information and Communication Technology, University of Phayao [Grant No. RMICT68006]

Acknowledgement

The author thanks the management and front-line officers of Chiangmuan Municipality, Phayao Province, for their cooperation throughout the two evaluation rounds, and the academic reviewers at the School of Information and Communication Technology, University of Phayao, who reviewed the content validity of the questionnaire instrument.

References

- [1] Wilbert O. Galitz, *The Essential Guide to User Interface Design, An Introduction to GUI Design Principles and Techniques*, John Wiley and Sons, vol. 29, no. 4, 1997. [Google Scholar] [Publisher Link]
- [2] Ananda Mohon Ghosh, Debashish Halder, and S.K. Alamgir Hossain, "Remote Health Monitoring System through IoT," 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV), Dhaka, Bangladesh, pp. 921-926, 2016. [CrossRef] [Google Scholar] [Publisher Link]

- [3] Sampada Kembhavi, Ravindra Gupta, and Gajendra Singh, "An Efficient Algorithm for Auto Upload and Chi-Square Test on Application Software," *International Journal of Advanced Computer Research*, vol. 3, no. 10, pp. 121-125, 2013. [[Google Scholar](#)] [[Publisher Link](#)]
- [4] Howard Lei et al., "A Statistical Analysis of the Effects of Scrum and Kanban on Software Development Projects," *Robotics and Computer-Integrated Manufacturing*, vol. 43, pp. 59-67, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [5] Afifa Lodhi, "Usability Heuristics as an Assessment Parameter: For Performing Usability Testing," *2010 2nd International Conference on Software Technology and Engineering*, San Juan, PR, USA, pp. V2-256-V2-259, 2010. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [6] Anand P. Mankodia, and Satish K. Shah, "Design of User Interface for Shot Boundary Detection using Classical Methods," *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and control Engineering*, vol. 1, no. 7, pp. 287-293, 2013. [[Google Scholar](#)] [[Publisher Link](#)]
- [7] Jakob Nielsen, "Enhancing the Explanatory Power of Usability Heuristics," *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp. 152-158, 1994. [[Google Scholar](#)]
- [8] Freddy Paz, and Jose Antonio Pow-Sang, "A Systematic Mapping Review of Usability Evaluation Methods for Software Development Process," *International Journal of Software Engineering and its Applications*, vol. 10, no. 1, pp. 165-178, 2016. [[Google Scholar](#)] [[Publisher Link](#)]
- [9] Daniela Quiñones, Cristian Rusu, and Virginica Rusu, "A Methodology to Develop Usability/user Experience Heuristics," *Computer Standards and Interfaces*, vol. 59, pp. 109-129, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [10] Wei-siong Tan, Dahai Liu, and Ram Bishu, "Web Evaluation: Heuristic Evaluation vs. user Testing," *International Journal of Industrial Ergonomics*, vol. 39, no. 4, pp. 621-627, 2009. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [11] Thankam Paul Thyvalikakath et al., "Comparative Study of Heuristic Evaluation and Usability Testing Methods," *Studies in Health Technology and Informatics*, vol. 143, pp. 322-327, 2009. [[Google Scholar](#)] [[Publisher Link](#)]
- [12] Enlie Wang, and Barrett Caldwell, "An Empirical Study of Usability Testing: Heuristic Evaluation vs. User Testing," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 46, no. 8, pp. 774-778, 2002. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [13] Yang Xu et al., "AI-Driven UX/UI Design: Empirical Research and Applications in FinTech," *Academia Nexus Journal*, vol. 3, no. 1, pp. 1-28, 2024. [[Google Scholar](#)] [[Publisher Link](#)]
- [14] Noptanit Chotisarn et al., "Usability Heuristic Evaluation for Phayao Province E-Government Portal, Thailand," *NU International Journal of Science*, vol. 13, no. 1, pp. 8-16, 2016. [[Google Scholar](#)]
- [15] Fillip Molodtsov, and Anastasija Nikiforova, "An Integrated Usability Framework for Evaluating Open Government Data Portals: Comparative Analysis of EU and GCC Countries," *Proceedings of the 25th Annual International Conference on Digital Government Research*, Association for Computing Machinery, New York, United States, pp. 899-908, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [16] Stefanie Di Russo, "Design and Taxes: A Case Study on Design Thinking in the Australian Taxation Office," *Proceedings of IASDR Interplay*, Brisbane, Australia, 2015. [[Google Scholar](#)]