https://doi.org/10.1016/j.rser.2021.111475

Development of a framework for evaluating the contents and usability of the building life cycle assessment tool

Minjin Kong [1], Minhyun Lee [2], Hyuna Kang [3], and Taehoon Hong [4]*

[1] Graduate Research Assistant, Department of Architecture and Architectural Engineering, Yonsei University, Seoul, Republic of Korea, E-mail: min920606@yonsei.ac.kr.

[2] Assistant Professor, Department of Building and Real Estate, The Hong Kong Polytechnic University, Kowloon, Hong Kong, E-mail: minhyun.lee@polyu.edu.hk.

[3] Graduate Research Assistant and Ph.D. Student, Department of Architecture and Architectural Engineering, Yonsei University, Seoul, Republic of Korea, E-mail: hyuna_kang@yonsei.ac.kr.

[4] Underwood Distinguished Professor, Department of Architecture and Architectural Engineering, Yonsei University, Seoul, Republic of Korea, E-mail: hong7@yonsei.ac.kr (CORRESPONDING AUTHOR).

* Yonsei University

50 Yonsei-ro, Seodaemun-gu, Seoul, 03722, Republic of Korea

82-2-2123-5788 (T); 82-2-365-4668 (F);

hong7@yonsei.ac.kr

Abstract

Over the past decades, various tools that can perform building life cycle assessment (LCA) as well as life cycle cost (LCC) or CO_2 analysis, have been developed. Even though these developed tools should be effectively evaluated and improved to encourage the continuous use of such tools, no research has been conducted on this matter. In this regard, this study sought to propose a framework for evaluating a building LCA tool from both the developer's and user's perspectives. In the developer evaluation process, experts evaluate if the design and implementation status (i.e., content status) are appropriate, and determine the design and implementation problems (i.e., content problems) based on six evaluation criteria, through content evaluation. In the user evaluation process, the users and evaluators determine the usability problems based on six usability attributes, through usability evaluation. The developer and user evaluation results are then interpreted through Satisfaction-Importance (S-I) analysis and Severity-Priority (S-P) analysis to prioritize the area of improvement and to determine the improvement strategy. To verify the proposed framework, a case study was conducted on an actual building LCA tool. The evaluation results showed that the problems corresponding to the assessment method and result should be preferentially improved in terms of content, while those corresponding to learnability, efficiency, and errors should be preferentially improved in terms of usability. Therefore, it is expected that the utilization of the proposed framework can effectively evaluate and improve various conventional building LCA tools in a reasonable way.

Keywords: *Building LCA tool; Usability evaluation; Content evaluation; Improvement strategy; Satisfaction-Importance analysis; Severity-Priority analysis.*

Abbreviation

BIM	*Building Information Modeling*
BRE	*Building Research Establishment*
CIMSCITY 2.0	*Carbon-Integrated construction Management System in CITY*
D&D	*Deconstruction and disposal*
HCI	*Human-Computer Interaction*
IEA	*International Energy Agency*
INDCs	*Intended Nationally Determined Contributions*
ISO	*International Organization for Standardization*
KIPFA	*Korea Internet Professional Association*
LCA	*Life cycle assessment*
LCC	*Life cycle cost*
$LCCO_2$	*Life cycle CO_2 emission*
O&M	*Operation and maintenance*
S-I	*Satisfaction-Importance*
S-P	*Severity-Priority*
UI	*User Interface*
UNEP	*United Nations Environmental Programme*

1. Introduction

The 2019 report jointly published by the United Nations Environment Programme (UNEP) and International Energy Agency (IEA) pointed out that buildings and the construction sector account for 39% of the global energy-related CO_2 emissions. To reduce the CO_2 emissions from building sector, most countries have included buildings and the construction sector in their Intended Nationally Determined Contributions (INDCs), a voluntary CO_2 emission reduction goal [1]. In a related move, the world has highlighted the critical importance not only of green building technologies that can reduce CO_2 emissions from buildings but also of technologies capable of evaluating the CO_2 emissions generated during a building's life cycle [2–5]. In response, various countries, companies, and research institutes have exerted efforts to develop building life cycle assessment (LCA) tools as a part of such technologies. Building LCA tools are designed to quantitatively evaluate various environmental impacts, including the CO_2 emissions that occur during the entire life cycle of a building. They have been used in many previous studies to analyze building materials and construction equipment that release large amounts of CO_2 into the atmosphere, and to determine how to reduce their CO_2 emissions [6–11]. IMPACT developed by Building Research Establishment (BRE) in the United Kingdom provides accurate life cycle cost (LCC), LCA, and life cycle CO_2 emission ($LCCO_2$) results for buildings based on the standardized specifications and dataset available in the United Kingdom. It can be also integrated into 3D modeling tools such as computer-aided design (CAD) or building information modeling (BIM) softwares not only for simple LCA applications but also for green building certification with Building Research Establishment Environmental Assessment Method (BREEAM) [12]. Similarly, the Tally application, a BIM software plugin jointly developed by KT Innovations, thinkstep, and Autodesk, can analyze the environmental impacts of an entire building through building material selection directly in an Autodesk Revit model [13]. It also allows LCA for green building certification with

Leadership in Energy and Environmental Design (LEED) and uses the life cycle inventory (LCI) database from GaBi software by thinkstep in Germany, which is the world's leading software for calculating the energy, cost, and environmental criteria based on the most current, comprehensive, and accurate LCA database [14,15]. Athena Impact Estimator for Buildings, developed in Canada, evaluates a whole building, both for new buildings and major renovations, based on LCA methodology, using the LCI database of the United States (U.S.) [16]. Both the Tally application and Athena Impact Estimator for Buildings are developed to be used for buildings specifically located in North America [17]. SimaPro is a product LCA tool developed by PRe Sustainability in the Netherlands that can be applied to the building sector and has the advantage of utilizing databases from various countries, such as the U.S., European Union countries, and Japan [18]. GEM-21P is a tool developed by Shimizu Corporation in Japan that calculates the CO_2 emissions of a building based on the building's statistical data and energy use per unit area [19,20].

The general purpose of building LCA tools is to effectively perform LCA on buildings and simultaneously provide the results to their users, and thus to support the users' decision-making and help in their communication with third parties [8,21]. Therefore, apart from developing a building LCA tool, it is necessary to evaluate if such a tool can actually achieve its purpose. In relation to the healthcare or education industry, much research has been done not only on the development of a healthcare or education tool to support the users' decision-making but also on the evaluation of the developed tool [22–28]. Milward [29] emphasized the need for the evaluation of such tools through research that showed that 95% of the existing healthcare tools are disengaged within a month due to problems with their contents or quality. Therefore, to increase the efficacy and sustainability of the developed tool, a rational evaluation of it needs to be conducted to determine if there are problems in its contents or quality (and if so, to identify these) and what should be done to achieve its purpose.

In the research area of building LCA tools, however, studies have been conducted mainly on the development and utilization of such tools [7,9,30,31] or on a comparative review of the developed building LCA tools [10,11,21], and a direct evaluation of the developed tools has rarely been conducted. In connection with the development and utilization of a building LCA tool, Ji et al. [7] used the building LCA tool based on case-based reasoning and a hybrid LCA to evaluate the environmental benefits of new educational facilities and to compare these with the existing cases. As the said study, however, focused only on the development, improvement, and utilization stages of the building LCA tool, it did not evaluate if the building LCA tool could successfully achieve its purpose. With respect to the comparative review of the developed building LCA tools, Haapio and Viitaniemi [21] classified and analyzed a variety of building LCA tools in accordance with criteria like type, purpose, and method, but claimed that evaluation through a comparison of the tools is very difficult and almost impossible because the purpose, method, and target vary depending on the building LCA tool. This claim, however, was limited to the comparison of various tools, and there was no comment on the evaluation of the building LCA tool itself. That is, an individual evaluation of a building LCA tool is required to increase its efficacy and sustainability and to ensure the proper dissemination of the LCA method.

To evaluate an individual building LCA tool, this study examined the previous research related to the evaluation of a healthcare tool. The purpose of the healthcare tool is similar to that of the building LCA tool, which is to support the users' decision-making by providing expert knowledge. According to the previous research on the measurement and evaluation of the quality of a mobile healthcare tool, various tools can be evaluated from both the developer's and user's perspectives [32]. From the developer's perspective, the contents of the developed tool are evaluated, and from the user's perspective, the usability of the tool is evaluated through its actual use.

- The following studies dealt with the evaluation of the contents of the developed tool from the developer's perspective. Stoyanov et al. [33] pointed out the absence of criteria for evaluating the information quality of mobile healthcare tools and carried out the content quality rating process. Garnett et al. [34] used the Delphi method to identify the essential functions for evaluating the contents of a mobile healthcare tool and evaluated the 12 essential functions selected by experts. The evaluation was performed by selecting criteria that could identify the contents of the tool based on the study.
- The following studies dealt with the evaluation of the usability of the developed tool from the user's perspective. Khajouei et al. [35] compared the heuristic evaluation and cognitive walkthrough results to identify a method that could effectively evaluate a healthcare tool and its users' interaction problems with it. Kang and Park [25] conducted a heuristic evaluation through five evaluators to determine the usability of mobile healthcare tool. Schnall et al. [36] also performed a usability assessment through five heuristic evaluators and ten end-user usability testers for the usability evaluation of the healthcare tool. These studies dealt with the evaluation of mobile healthcare tools from the user's perspective.
- The following study dealt with the evaluation of the developed tool from both the developer's and user's perspectives. Attwood et al. [37] evaluated the tool's contents by confirming the presence of behavior change techniques for the developed mobile healthcare tool and then identified the users' opinions about such tool's acceptability, usability, and potential effectiveness through qualitative interviews.

As stated above, studies related to the evaluation of healthcare tools have been conducted from either the developer's or user's perspective, and there is a dearth of research considering both the developer's and user's perspectives. Lew et al. [38] emphasized that both the information quality and actual usability of the tool should be evaluated, and suggested a method for evaluating information quality in the existing International Organization for

Standardization (ISO) 25010 standard. Therefore, to increase the efficacy and sustainability of the developed tools, evaluation should not be performed in terms only of either their contents or usability, but both their contents and usability should be critically considered in the evaluation process [22,24].

This study sought to propose a framework that could be applied to the evaluation of building LCA tools. The framework aims for building LCA tool developers who want to increase the efficacy and sustainability of their building LCA tool, providing them with improvement strategies for quality assurance and control. The framework was made to consist of (i) an evaluation procedure to evaluate the building LCA tool from both the developer's and user's perspectives, and (ii) an interpretation procedure to deduce improvement strategies for the building LCA tool by analyzing and interpreting the evaluation results. According to the previous research of evaluating healthcare tool mentioned above, experts in the corresponding field evaluated its contents from the developer's perspective while usability experts or general users of healthcare tool evaluated its usability from the user's perspective. Referring to this, building LCA, LCC, or $LCCO_2$ experts are selected to evaluate the contents of a building LCA tool while usability experts and general users are selected to evaluate its usability in the proposed framework. To verify the effectiveness of the framework, the proposed framework was applied to Carbon-integrated Construction Management System in CITY (CIMSCITY 2.0), a building LCA tool developed by this research team. Then evaluation and interpretation procedures were carried out to deduce the improvement strategy that would increase the accessibility and sustainability of CIMSCITY 2.0. The framework can be applied to a variety of existing building LCA tools, whose sustainability and accessibility this study intended to increase, thereby contributing to the dissemination of building LCA tools, which are essential for reducing the various environmental impacts of construction, including CO_2 emissions.

2. Materials and Methods

The framework for evaluating building LCA tools proposed in this study is divided mainly into two parts: (i) evaluation procedure and method; and (ii) interpretation procedure and method. The procedure for and method of evaluating a building LCA tool from both the developer's and user's perspectives are presented in section 2.1, and the procedure for and method of effectively improving a building LCA tool based on its evaluation results are presented in section 2.2 (refer to Fig. 1).

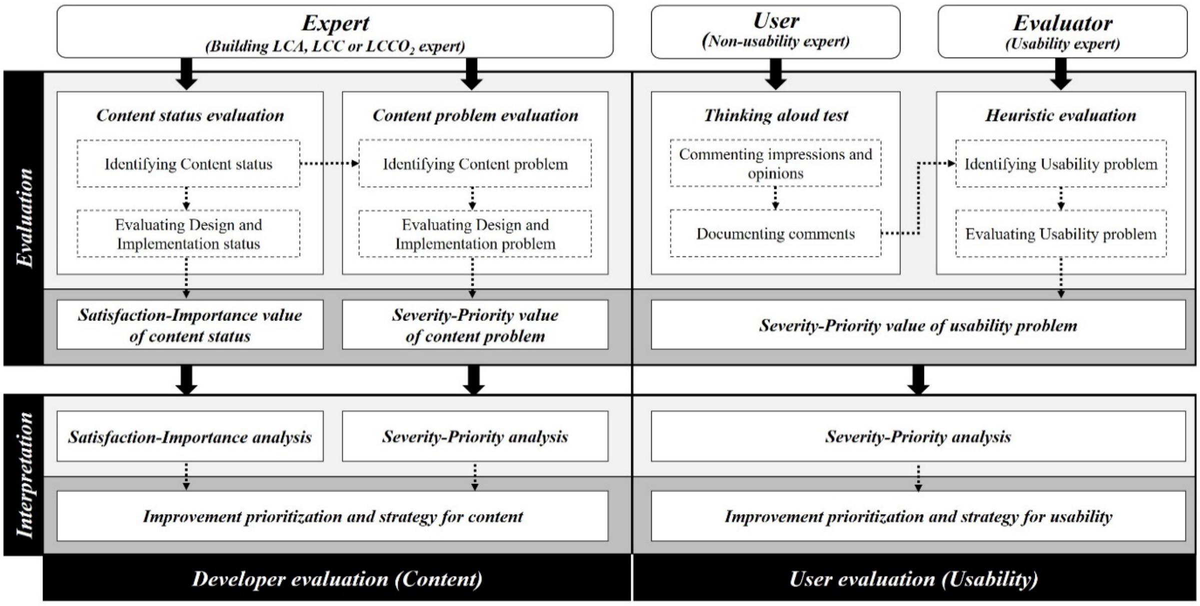


Fig. 1. Framework for evaluating the building LCA tool

2.1. Evaluation procedure and method for the building LCA tool

To effectively examine the status and problems of a building LCA tool from both the developer's and user's perspectives, evaluation of the building LCA tool is performed from two perspectives (refer to Fig. 2): (i) developer evaluation: content evaluation based on six evaluation criteria; and (ii) user evaluation: usability evaluation based on six usability attributes.

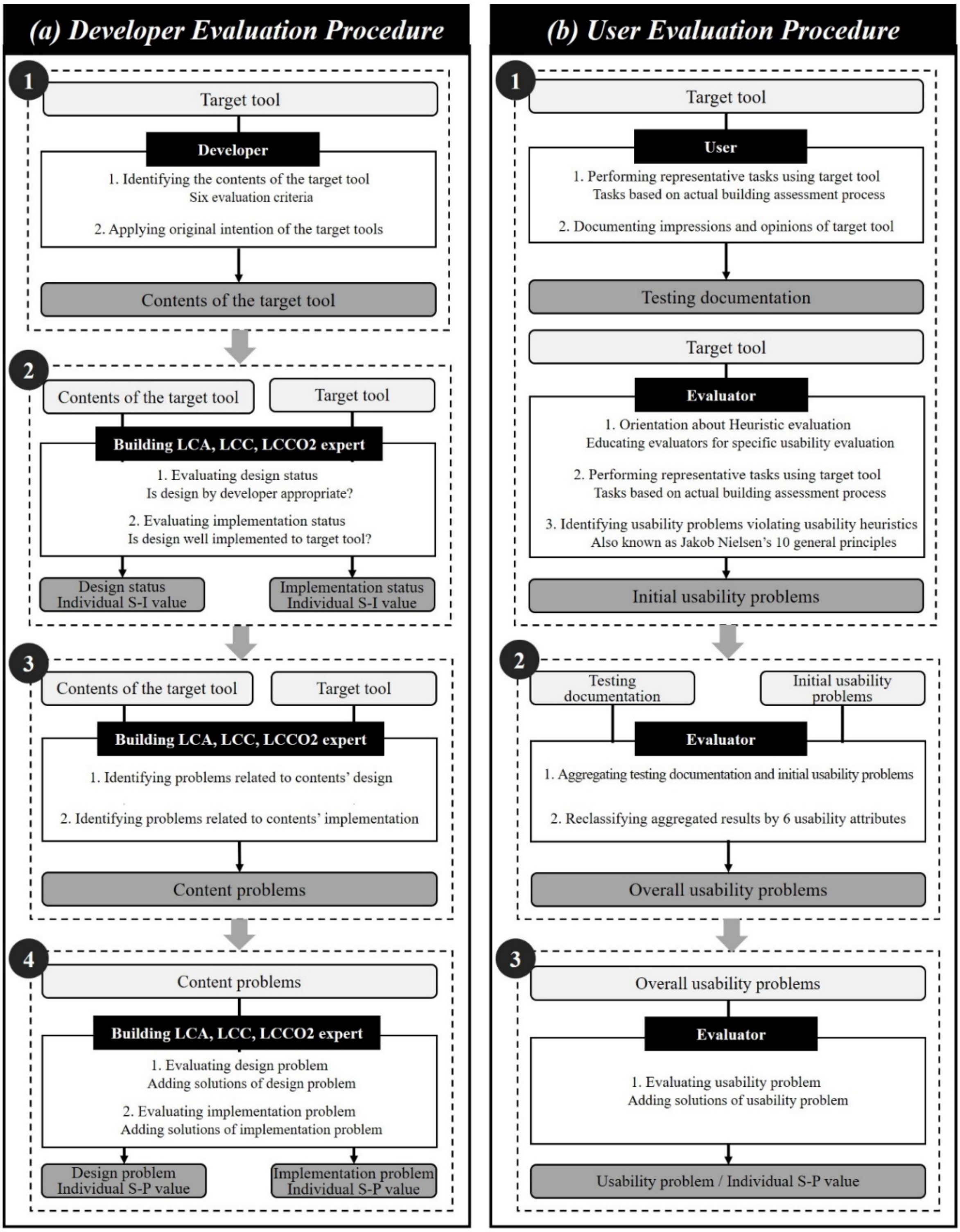


Fig. 2. (a) Developer evaluation procedure; and (b) User evaluation procedure

2.1.1. Developer evaluation procedure and method

The items for evaluating the contents of a building LCA tool should be defined before proceeding with the evaluation from the developer's perspective. To make the proposed framework applicable to various building LCA tools, evaluation criteria were derived by referring to studies that systematically classified and analyzed the existing building LCA tools [6,10,21,39,40]. As a result, six evaluation criteria for classifying and identifying the contents of building LCA tools were defined, as summarized in Table 1.

Table 1. Six evaluation criteria for identifying the contents of the building LCA tool

Evaluation Criteria		**Description**
Criterion 1	"Who" the tool is made for	User of the tool
Criterion 2	"What" the tool analyzes	Target of the tool
Criterion 3	"Why" the tool is needed	Specific purpose of the tool
Criterion 4	"How" the tool analyzes	Assessment method
Criterion 5	"Result" of the tool	Assessment result
Criterion 6	"Reference" of the tool	Assessment reference

Criterion 1 is an evaluation criterion for evaluating to "whom" the tool gives the analysis results. Criterion 2 is an evaluation criterion for evaluating "what" is analyzed by the tool. Criterion 3 is an evaluation criterion for evaluating "why" the tool is needed and what its specific purpose is. Criterion 4 is an evaluation criterion for evaluating "how" the tool performs the evaluation. Criterion 5 is an evaluation criterion for evaluating what analysis "results" the tool provides. Lastly, Criterion 6 is an evaluation criterion for evaluating the "reference" used in the tool.

Based on the six evaluation criteria, the developer evaluation is carried out in four steps (refer to Fig. 2 (a)): (i) Step 1: Identifying the contents by evaluation criteria; (ii) Step 2: Performing content status evaluation; (iii) Step 3: Identifying the content problems; and (iv) Step 4: Performing content problem evaluation. The results obtained from steps 2 and 4 are used in the interpretation procedure after the evaluation procedure.

2.1.1.1. Identifying the contents by evaluation criteria

To evaluate the contents of a building LCA tool from the developer's perspective, the contents of the target building LCA tool (in short, the target tool) for evaluation should be identified in accordance with the six evaluation criteria defined earlier. In principle, the developer of the target tool should directly identify the contents of the target tool corresponding to the evaluation criteria. Also, it is essential to define the contents of the target tool in as detailed and specific a way as possible because the evaluator performs the actual developer evaluation based on the results in the next step.

2.1.1.2. Performing content status evaluation

To evaluate the contents of the target tool identified in accordance with the six evaluation criteria, experts related to the target tool (e.g., building LCA, LCC, or $LCCO_2$ experts) participate in the process. For a meaningful qualitative evaluation, it is recommended that five or more experts participate in the process [41]. Each expert must have direct experience with the target tool, and separately performs content status evaluation based on the contents of the target tool identified in accordance with the six evaluation criteria. In the content status evaluation, the expert evaluates the opinions on the two types of content status (the design and implementation status) using the satisfaction-importance value (S-I value). The satisfaction value is a score indicating how satisfactory the content status is while the importance value is a score indicating how important the content status is. Each of these values uses a Likert scale from 1 (very low) to 5 (very high).

- Evaluating the design status of the contents: Whether the contents of the target tool identified in accordance with the six evaluation criteria were properly designed to achieve the general purpose of the target tool from the expert's perspective is evaluated. The expert presents the design status of the contents of the target tool for the six evaluation criteria, from Criterion 1 to 6, as the S-I value.

- Evaluating the implementation status of the contents: Whether the contents of the target tool identified in accordance with the six evaluation criteria were properly implemented in the actual target tool from the expert's perspective is evaluated. The expert presents the implementation status of the contents of the target tool for the six evaluation criteria, from Criterion 1 to 6, as the S-I value.

2.1.1.3. Identifying the content problems

Each expert identifies all the parts that he thinks has problems with the content status of the target tool evaluated based on the six evaluation criteria. Like the content status, the content problems are also divided into two types for evaluation: design and implementation problems. Where the problem occurs, a specific explanation of the problem and the reason for the occurrence of the problem are presented for each problem.

- Identifying the design problem: The problem corresponding to the design status is identified if the evaluation reveals that the contents of the target tool identified in accordance with the six evaluation criteria were not properly designed to achieve the general purpose of the target tool from the expert's perspective.
- Identifying the implementation problem: The problem corresponding to the implementation status is identified if the evaluation reveals that the contents of the target tool identified in accordance with the six evaluation criteria were not properly implemented in the actual target tool from the expert's perspective.

2.1.1.4. Performing content problem evaluation

Based on the content problem identified earlier, each expert individually performs content problem evaluation. In the content problem evaluation, the expert evaluates the opinions on two types of content problem (i.e., design and implementation problems) using the S-P value. The severity value is a score indicating how difficult it is to solve the content problem while the priority value is a score indicating how preferentially the content problem needs to be improved. Each value uses a Likert scale from 1 (very easy or very low) to 5 (very difficult or very high). Lastly, each expert provides solutions for each content problem.

2.1.2. User evaluation procedure and method

Usability, a key factor for the successful implementation of web applications, systems, and tools, is a quality factor for identifying their ease-of-use. It should be evaluated using a method suitable for the evaluation purpose of the building LCA tool among the various usability evaluation methods [42,43]. According to the systematic mapping study conducted by Fernandez et al. [42], usability evaluation has been performed in the implementation phase of a web application or tool in a total of 315 studies, 115 (36.5%) of which used testing as the usability evaluation method, 78 (24.8%) of which used inspection, and 68 (21.6%) of which used inquiry. Except for inquiry, which was used as a supplementary item, usability evaluation methods with different characteristics, such as testing and inspection, were used in the usability evaluation of the building LCA tool.

- Testing: While users who are non-usability experts are performing typical tasks with the target tool, an inspector observes them and evaluates the tool's usability when the thinking aloud test, the most representative method of verbalizing the users' opinions about the target tool among the various subtypes of testing, is used [44,45]. In the thinking aloud test, the target tool is evaluated through the user's response to the user interface (UI). Jakob Nielsen, the world's leading expert on usability and an authoritative usability consultant, recommends that to proceed with the thinking aloud test, five users participate [46]. The analysis results of more than 80 usability consulting projects also suggest that the suitable number of participants is five [47].

- Inspection: Evaluators who are usability experts evaluate the usability of the target tool based on certain criteria while using the target tool when heuristic evaluation, the most frequently used among the various subtypes of inspection, is used [35]. Heuristic evaluation is a usability evaluation method proposed by Jakob Nielsen, and is a method for human-computer interaction (HCI) experts to discover and evaluate problems that violate the heuristic principle. The optimal number of participants in terms of cost effectiveness is three to five because the participation of five evaluators can lead to a verified evaluation result level, and the ratio of cost to problem discovery efficiency decreases rapidly when the number exceeds five [48,49].

According to the existing research, non-usability expert (user) based evaluation and usability expert (evaluator) based evaluation cannot be replaced, but they can be used complementarily [50]. Therefore, a comprehensive usability evaluation of the target tool was performed based on the thinking aloud test through user participation, and heuristic evaluation was performed through evaluator participation. Here, the user who performs the thinking aloud test refers to the user of the tool indicated in Criterion 1 from the six evaluation criteria (refer to Table 1). Conversely, the evaluator who performs the heuristic evaluation refers to a usability expert who has no expertise in building LCA. The differences between the thinking aloud test and heuristic evaluation, however, were not dealt with in this study.

Based on the usability evaluation method selected earlier, user evaluation is performed in three steps (refer to Fig. 2 (b)): (i) Step 1: Identifying the usability problems; (ii) Step 2: Reclassifying usability problems; and (iii) Step 3: Performing usability problem evaluation. The results obtained from step 3 are used in the interpretation procedure after the evaluation procedure.

2.1.2.1. Identifying the usability problems

In the thinking aloud test, all the users' opinions expressed while they are using the target tool are recorded by the inspector. As the users, who are non-usability experts, cannot precisely analyze the problem themselves, the recorded comments are delivered to the evaluators, who are usability experts. Before any evaluation by the evaluators, a detailed manual for operating the target tool was provided, as the evaluators have no expertise in building LCA as mentioned above. Based on the provided manual, the evaluators identify the initial usability problems that violate the 10 usability heuristics, also known as Jakob Nielsen's 10 general principles for interaction design (refer to Table 2) [51].

Table 2. Jakob Nielsen's 10 general principles for interaction design [51]

Usability Heuristics	Definition
Visibility of system status	The system should indicate its current state to the users.
Match between system and the real world	The system should speak a language familiar to the users.
User control and freedom	The users should be able to cancel or redo the process of the system.
Consistency and standards	The system should be consistent to prevent misunderstanding of the users.
Error prevention	Errors should be eliminated; the users should be able to check actions that might provoke errors
Recognition rather than recall	The users should recognize the process or functions of the system not recall.
Flexibility and efficiency of use	Both novice and expert users should be able to use the system efficiently.
Aesthetic and minimalist design	Irrelevant information should be deleted and minimalized.
Help users recognize, diagnose, and recover from errors	If the error occurs, the users should be able to understand the reason why and way to recover the error.
Help and documentation	Manual or sample of the system should be prepared to help and assist the users.

2.1.2.2. Reclassifying the usability problems

The users' opinions expressed in the thinking aloud test and the usability problems discovered by the evaluators through heuristic evaluation are combined to reclassify the usability problems according to six usability attributes. Table 3 shows a summary of the six usability attributes defined by ISO and Jakob Nielsen [23,35,52,53]. When usability problem A derived from the thinking aloud test and the heuristic evaluation of the target tool is judged to undermine the learnability of the target tool, it is classified as a learnability-related usability problem.

Table 3. Six usability attributes

Usability Attribute	Definition	Reference
Effectiveness	How accurately can users achieve specified goals by target tool?	ISO 9241 [54]
Efficiency	How much resource (time, effort) are required to achieve specified goals by target tool?	ISO 9241 [54]
Satisfaction	How fun is it to use target tool?	ISO 9241 [54]
Learnability	How easily can first-time users achieve specified goals by target tool?	Nielsen 1993 [52]
Memorability	How proficiently can users operate target tool if user disused target tool long ago?	Nielsen 1993 [52]
Errors	How many errors can users occur? How severe are the errors, and can they be recovered?	Nielsen 1993 [52]

2.1.2.3. Performing usability problem evaluation

Based on the usability problems reclassified above, each evaluator individually performs usability problem evaluation. In the usability problem evaluation, the evaluator evaluates the opinions about the usability problems using the S-P value. As mentioned earlier, the severity value is a score indicating how difficult it is to solve the usability problem while the priority value is a score indicating how preferentially the usability problem needs to be improved. Each value uses a Likert scale from 1 (very easy or very low) to 5 (very difficult or very high). Lastly, each evaluator provides solutions for each usability problem.

2.2. Interpretation procedure and method for the building LCA tool

To effectively improve the target building LCA tool based on its evaluation results, it is analyzed from two perspectives (refer to Fig. 3): (i) Satisfaction-Importance (S-I) analysis and improvement prioritization for content status evaluation; and (ii) Severity-Priority (S-P) analysis and improvement prioritization for content and usability problem evaluation.

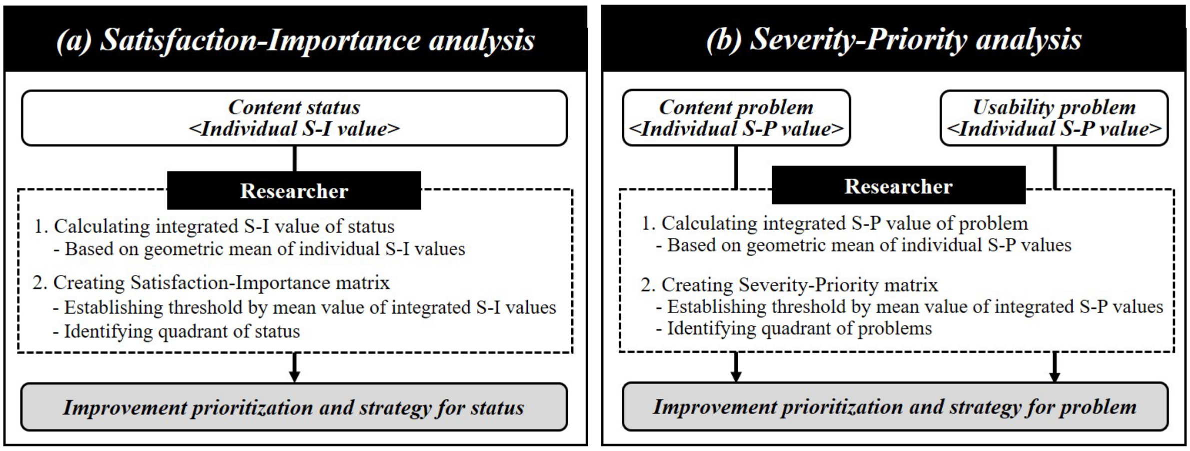


Fig. 3. (a) Satisfaction-Importance (S-I) analysis procedure; and (b) Severity-Priority (S-P) analysis procedure

2.2.1. Satisfaction-Importance (S-I) analysis and improvement prioritization

The S-I value, the content status evaluation result obtained from the developer evaluation, is interpreted through S-I analysis (refer to Fig. 3 (a)). In S-I analysis, the performance in the existing importance-performance analysis is analyzed by replacing it with the expert's satisfaction based on importance-performance analysis, a methodology for analyzing which of the elements making up a certain business or product is most noticeable [55–58]. Based on the matrix composed of the satisfaction degree on the x-axis and the importance degree on the y-axis (i.e., S-I matrix, refer to Fig. 4 (a)), the distribution of the S-I value on the content status of the target tool is analyzed [59]. In this case, the integrated S-I value for content status H can be calculated using the geometric mean of the individual S-I values assigned by n number of experts to content status H (refer to Eqs. (1) and (2)). To relatively evaluate the integrated S-I values for each content status using the S-I matrix, the said matrix should be divided into four quadrants based on the threshold obtained from the calculation of the S-I threshold when the said threshold can be calculated using the mean value of the S-I values for all the m number of content status (refer to Eqs. (3) and (4)).

$$b_H = [\prod_{I=1}^{n} b_{HI}]^{1/n} = \sqrt[1/n]{b_{H1} \times b_{H2} \cdots b_{Hn}} \quad (1)$$

$$c_H = [\prod_{I=1}^{n} c_{HI}]^{1/n} = \sqrt[1/n]{c_{H1} \times c_{H2} \cdots c_{Hn}} \quad (2)$$

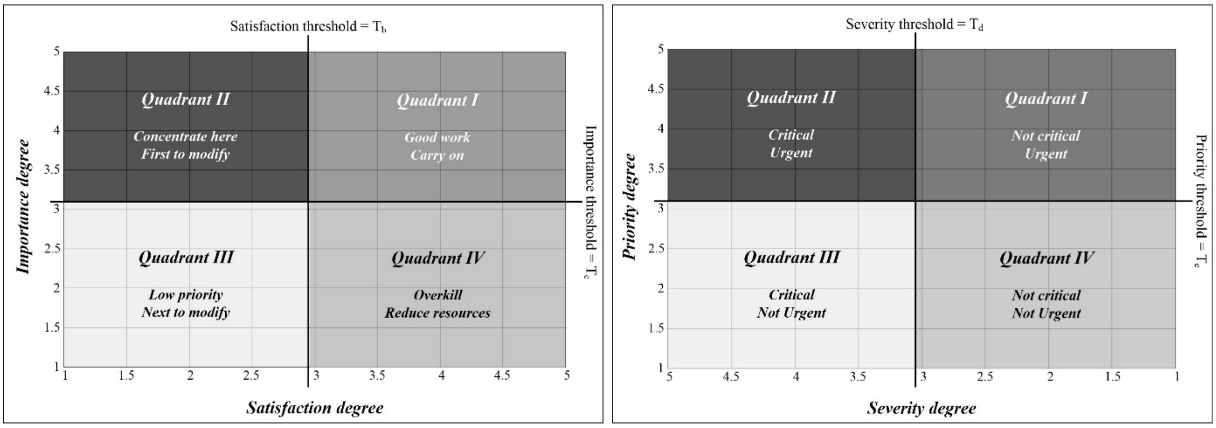
$$T_b = \sum_{H=1}^{m} b_H / m \quad (3)$$

$$T_c = \sum_{H=1}^{m} c_H / m \quad (4)$$

where b_H stands for satisfaction value of content status H, b_{HI} stands for individual satisfaction value of content status H by expert I, c_H stands for importance value of content status H, c_{HI} stands for individual importance value of content status H by expert I, T_b stands for satisfaction threshold, T_c stands for importance threshold, n is number of experts, and m is number of content status.

The S-I matrix is divided into four quadrants based on the S-I threshold, and each content status represented in the S-I matrix has different meanings in terms of satisfaction and importance depending on the quadrant where it is included. The content status included in the first quadrant (Quadrant I) corresponds to the strong point of the target tool because both its satisfaction and importance are high, so there is no need for improvement. The content status included in the second quadrant (Quadrant II) corresponds to the weak point of the target tool because its importance is high while its satisfaction is low, so improvement needs to be made most preferentially. The content status included in the third quadrant (Quadrant III) has lower improvement priority than that included in the second quadrant because both its satisfaction and importance are low. Lastly, in the case of the content status included in the fourth quadrant

(Quadrant IV), its satisfaction is excessively high compared to its low importance. Therefore, the resources required for the content status included in the fourth quadrant should be reduced to invest in the improvement of the content status included in the second quadrant (refer to Fig. 4 (a)). As such, the improvement priority of the content status of the target tool can be determined through S-I analysis, as follows: (i) the first improvement priority: the content status in Quadrant II; and (ii) the second improvement priority: the content status in Quadrant III. As the content status in Quadrants I and IV achieve a degree of satisfaction above a certain level, they are regarded as items requiring no improvement.



(a) Satisfaction-Importance matrix *(b) Severity-Priority matrix*

Fig. 4. (a) Satisfaction-Importance (S-I) matrix; and (b) Severity-Priority (S-P) matrix

2.2.2. Severity-Priority (S-P) analysis and improvement prioritization

The S-P value of the content problems obtained from the developer evaluation and the S-P value of the usability problems obtained from the user evaluation are interpreted through S-P analysis (refer to Fig. 3 (b)). In S-P analysis, prioritization of the problems of the target tool instead of the bug in the existing research is performed based on the research of Zhang and Versteeg, which performed prioritization of various bugs for effective software program bug fixing [60]. Based on the matrix composed of the severity degree on the x-axis and the priority degree on the y-axis (i.e., S-P matrix, refer to Fig. 4 (b)), the distribution of the S-P value is interpreted with respect to the content or usability problem of the target tool. In this case, the integrated S-P value for content problem J can be calculated using the geometric mean of the individual S-P values assigned by n number of experts or evaluators to problem J, respectively (refer to Eqs. (5) and (6)). To relatively evaluate the integrated S-P values for each problem using the S-P matrix, the said matrix should be divided into four quadrants based on the threshold obtained from the calculation of the S-P threshold when the S-P threshold can be calculated using the S-P value for all the l number of contents or usability problems (refer to Eqs. (7) and (8)).

$$d_J = \left[\prod_{K=1}^{n} d_{JK}\right]^{1/n} = \sqrt[1/n]{d_{J1} \times d_{J2} \cdots d_{Jn}} \tag{5}$$

$$e_J = \left[\prod_{K=1}^{n} e_{JK}\right]^{1/n} = \sqrt[1/n]{e_{J1} \times e_{J2} \cdots e_{Jn}} \tag{6}$$

$$T_d = \sum_{J=1}^{l} d_J / l \tag{7}$$

$$T_e = \sum_{J=1}^{l} e_J / l \tag{8}$$

where d_J stands for severity value of problem J, d_{JK} stands for individual severity value of problem J by evaluator K, e_J stands for priority value of problem J, e_{JK} stands for individual priority value of problem J by evaluator K, T_d stands for severity threshold, T_e stands for priority threshold, and n is number of experts or evaluators, and l is number of problems.

The S-P matrix is divided into four quadrants based on the S-P threshold calculated in this way, and each problem represented in the S-P matrix has different meanings in terms of severity and priority, depending on the quadrants where it is included. The problem included in the first quadrant (Quadrant I) can achieve great effects even with a small investment in resources because its priority is high while its severity is low. The problem included in the second quadrant (Quadrant II) is a top priority for improvement but requires a huge amount of resources because of its high severity and priority. The problem included in the third quadrant

(Quadrant III) also requires a large amount of resources for improvement due to its higher severity compared to its priority, but its low need for improvement decreases the efficiency of improvement. Lastly, the problem included in the fourth quadrant (Quadrant IV) can show improvement with a small investment in resources due to its low severity and priority, but the need for improvement is low (refer to Fig. 4 (b)). As such, the improvement priority of the content problem of the target tool can be determined based on S-P analysis, as follows: (i) the first improvement priority: the problems in Quadrant II; (ii) the second improvement priority: the problems in Quadrant I; (iii) the third improvement priority: the problems in Quadrant IV; and (iv) the fourth improvement priority: the problems in Quadrant III. First, although the problems in Quadrants I and II need urgent improvement, the top priority for improvement should be given to the problems in Quadrant II as they require more resources, such as time and money. On the other hand, the problems in Quadrant I become the second priority because they can be managed more flexibly compared to the problems in Quadrant II. Meanwhile, the problems in Quadrants III and IV do not require urgent improvement, and their need for improvement is also low, but the problems in Quadrant IV can show immediate improvement by investing in surplus resources as they can be improved with only a small amount of resources. On the other hand, the content problems in Quadrant III become the last improvement priority as their improvement effects are expected to be insignificant compared to the amounts of resources needed for improvement.

3. Case Study

To verify the effectiveness of the proposed framework, a case study was performed by applying the proposed framework to CIMSCITY 2.0, a building LCA tool developed by this research team. CIMSCITY 2.0 is a web-based building LCA tool that can effectively perform assessment related to buildings' CO_2 emissions as well as the environmental impacts and costs in terms of the whole life cycle of the building, and the utilization of the analysis results obtained using the tool makes it possible to support the users' decision-making and to help them communicate with third parties.

CIMSCITY 2.0 contains a total of four stages: (i) Design stage; (ii) Construction stage; (iii) Operation and maintenance (O&M) stage; and (iv) Deconstruction and disposal (D&D) stage. None of these stages require the use of a BIM model for the simplicity of the tool. In each stage, LCA, LCC analysis, and even $LCCO_2$ analysis of buildings are performed using a variety of methods as follows.

- Design stage: A model has been developed and presented to predict the life cycle cost, energy, or CO_2 emissions of the target building in the design stage based on the data of existing buildings. Here, an advanced-Case Based Reasoning (A-CBR) is used to develop the prediction model. The contents of CIMSCITY 2.0 **-** Design stage are presented in Table 4 according to the proposed six evaluation criteria [61].

Table 4. Contents of CIMSCITY 2.0 - Design stage

Criteria	Contents	
Criterion 1 User of the tool	➢ Building design experts including architects, designers, engineers, and contractors ➢ LCC or LCA experts including researchers and consultants ➢ Investors and building owners with expert assistance	
Criterion 2 Target of the tool	➢ Buildings in the design stage with known design variable	
Criterion 3 Purpose of the tool	➢ To predict the life cycle cost, energy, or CO_2 emissions of the target building before it is built ➢ To retrieve similar cases of the target building in the design stage for benchmarking	
Criterion 4 Assessment method	Input data	➢ Database (i.e., independent and dependent variables) on case buildings ➢ Predictive values for the dependent variable of case buildings by multiple regression analysis (MRA) and artificial neural network (ANN) ➢ Information (i.e., independent variables) on the target building
	Analysis method	➢ Case Based Reasoning (CBR) ➢ Advanced-Case Based Reasoning (A-CBR) [1)]
Criterion 5 Assessment result	➢ Training results: (i) Prediction performance of prediction models; and (ii) similar cases of training dataset ➢ Case testing results: (i) Predictive value for the dependent variable of the target building; and (ii) similar cases of target building	
Criterion 6 Assessment reference	1) A-CBR: A CBR model with improved prediction accuracy by applying Minimum Criterion for scoring the Attribute Similarity (MCAS), Range of the Attribute Weight), RCS (Range of the Case Selection (RAW), and Tolerance Range of the Cross-Range between MRA and ANN (TRCRMA) [61]	

- Construction stage: A model has been developed and presented to monitor and estimate the cost and CO_2 emissions of the target building during the construction stage, from material manufacturing to site construction. The planned and actual cost and CO_2 emissions of the target building during the construction stage has been compared based on the project schedule using an earned value management system (EVMS). The contents of CIMSCITY 2.0 - Construction stage are presented in Table 5 according to the proposed six evaluation criteria [9].

Table 5. Contents of CIMSCITY 2.0 - Construction stage

Criteria	Contents	
Criterion 1 User of the tool	➢ Building construction experts including architects, designers, engineers, and contractors ➢ LCC or LCA experts including researchers and consultants ➢ Investors and building owners with expert assistance	
Criterion 2 Target of the tool	➢ Reinforced concrete structured buildings ➢ Steel structured buildings	
Criterion 3 Purpose of the tool	➢ To monitor and estimate the cost and CO_2 emissions of the target building during the construction stage based on the project schedule ➢ To compare the planned and actual cost and CO_2 emissions of the target building during the construction stage and to bridge the gap between them	
Criterion 4 Assessment method	Input data	➢ Material quantity ➢ Transportation equipment and distance ➢ Construction equipment ➢ Project schedule
	Analysis method	➢ LCA 1) ➢ Earned value management system (EVMS) 2)
Criterion 5 Assessment result	➢ The cost and CO_2 emissions of the target building based on the planned project schedule ➢ The cost and CO_2 emissions of the target building based on the actual project schedule	
Criterion 6 Assessment reference	1) LCA database: National LCI database, heat values of various fuels, carbon emission factors, and electricity emission factors 2) EVMS: A project management method for measuring project performance and progress using the planned and earned value based on the cost and schedule - Earned value curve: Standard construction cost distribution curve	

• O&M stage: A model has been developed and presented to estimate the life cycle cost and environmental impact of the target building during the O&M stage based on the Bill of Quantities considering both building materials and operational energy. The contents of CIMSCITY 2.0 - O&M stage are presented in Table 6 according to the proposed six evaluation criteria [30].

Table 6. Contents of CIMSCITY 2.0 - O&M stage

Criteria	Contents	
Criterion 1 User of the tool	➢ Building O&M experts including architects, designers, engineers, and contractors ➢ LCC or LCA experts including researchers and consultants ➢ Investors and building owners with expert assistance	
Criterion 2 Target of the tool	➢ Buildings in the O&M stage with Bill of Quantities	
Criterion 3 Purpose of the tool	➢ To estimate the life cycle cost and environmental impact of the target building in the O&M phase based on Bill of Quantities ➢ To calculate the life cycle cost of the target building considering the repair and replacement of major building materials and energy consumption during the O&M stage ➢ To calculate the life cycle environmental impact of the target building considering both the direct and indirect energy consumption from materials and energies	
Criterion 4 Assessment method	Input data	➢ LCC analysis assumptions: (i) Analysis period; (ii) real discount rate; and (iii) inflation rate ➢ Bill of Quantities ➢ Repair and replacement rate of major building materials ➢ Energy (i.e., electricity and gas) consumption
	Analysis method	➢ LCC analysis (Present worth method) 1) ➢ LCA 2)
Criterion 5 Assessment result	➢ Life cycle cost of the target building ➢ Life cycle environmental impact of the target building ➢ Life cycle environmental cost of the target building	
Criterion 6 Assessment reference	1) LCC database: National consumer price index, nominal interest rate, electricity rates increase rate, gas rates increase rate, certified emission reduction price, and repair and replacement rate 2) LCA database: National LCI database and national Input-Output table	

• D&D stage: A model has been developed and presented to estimate the environmental impact during the D&D stage of the target building based on the Bill of Quantities and transportation distance. The contents of CIMSCITY 2.0 - D&D stage are presented in Table 7 according to the proposed six evaluation criteria.

Table 7. Contents of CIMSCITY 2.0 - D&D stage

Criteria	Contents	
Criterion 1 User of the tool	➢ Building D&D experts including architects, designers, engineers, and contractors ➢ LCC or LCA experts including researchers and consultants ➢ Investors and building owners with expert assistance	
Criterion 2 Target of the tool	➢ Reinforced concrete structured buildings ➢ Steel structured buildings ➢ Steel reinforced concrete structured buildings	
Criterion 3 Purpose of the tool	➢ To estimate the environmental impact generated from all energy and waste during the D&D stage of the target building based on Bill of Quantities ➢ To calculate the environmental impact during the D&D stage considering both the direct and indirect energy consumption of electricity and fossil energy ➢ To calculate the environmental impact during the D&D stage considering transportation distance according to the location of the target building	
Criterion 4 Assessment method	Input data	➢ Building location ➢ Building and structure type ➢ Bill of Quantities ➢ Deconstruction and transportation equipment
	Analysis method	➢ LCA [1)]
Criterion 5 Assessment result	➢ Direct and indirect energy consumption during the D&D stage ➢ CO_2 emissions during incineration ➢ Environmental impact during the D&D stage	
Criterion 6 Assessment reference	[1)] LCA database: National LCI database and national Input-Output table	

The web-based UI of CIMSCITY 2.0 consists of six main pages: (a) Login page; (b) Main page; (c) Design stage page; (d) Construction stage page; (e) O&M stage page; and (f) D&D stage page (refer to Fig. 5).

- Login page: Enter ID and password to login and experience CIMSCITY 2.0.
- Main page: Select one of the four stages (i.e., design, construction, O&M, or D&D stage) based on the intention of the user.
- Design stage page: A cost, energy, or CO_2 prediction model in the building design stage consists of four steps: (i) Data setting; (ii) model establishment; (iii) training result (including both prediction performance of prediction models and similar cases of training dataset); and (iv) case testing.
- Construction stage page: A cost, CO_2 and schedule management model in the building construction stage consists of five steps: (i) Material manufacturing; (ii) material

transportation; (iii) site construction; (iv) schedule information; and (v) EVMS result.

- O&M stage page: A life cycle cost and environmental impact evaluation model in the building O&M stage consists of five steps: (i) Basic information; (ii) construction information; (iii) O&M information; (iv) D&D information; and (v) evaluation result.
- D&D stage page: A cost and environmental impact evaluation model in the building D&D stage consists of four steps: (i) Basic information; (ii) material quantity input; (iii) waste quantity calculation; and (iv) evaluation result.

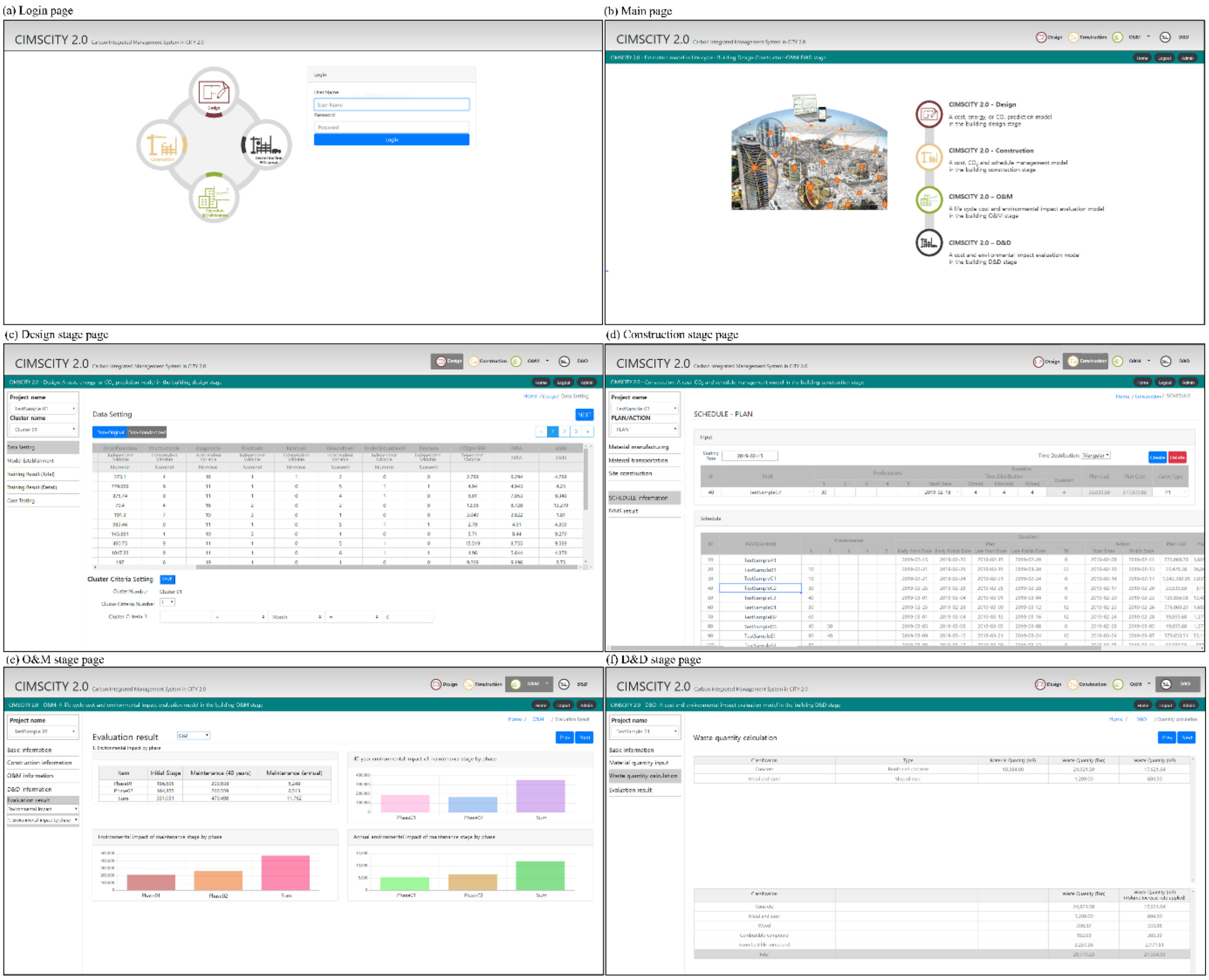


Fig. 5. Screenshots of CIMSCITY 2.0

4. Results and Discussions

In this study, both the developer and user evaluations of CIMSCITY 2.0 developed by this research team were performed by applying the proposed framework to an actual building LCA tool for a case study. Five experts participated in the developer evaluation while five users for the thinking aloud test and five evaluators for heuristic evaluation took part in the user evaluation. All the evaluations by the experts, users, and evaluators were conducted after approval by the Institutional Review Board (IRB) of Yonsei University in South Korea [IRB No. 7001988-201910-HR-707-03].

- Developer evaluation of CIMSCITY 2.0: A total of five building LCA, LCC, or $LCCO_2$ experts with a Ph.D. in Architectural Engineering who agreed to participate in the evaluation process were recruited for the developer evaluation. The developer identified the contents of CIMSCITY 2.0 for the four stages based on the six evaluation criteria (refer to Tables 4 to 7), and then delivered the completed contents of CIMSCITY 2.0 for each stage to five experts. The five experts evaluated the contents of CIMSCITY 2.0 in accordance with the developer evaluation procedure. The five experts first evaluate the design status of CIMSCITY 2.0 for each stage starting from Criterion 1 by answering the question "Is defining users of the target tool as building experts, LCC or LCA experts, or investors and building owners with expert assistance appropriate to achieve the general purpose of the target tool?". Then, its implementation status for Criterion 1 is evaluated by answering the question "After using the target tool, is the design of the target tool actually well implemented for the defined users (i.e., building experts, LCC or LCA experts, or investors and building owners with expert assistance)?". After evaluating satisfaction and importance of the design and implementation status in the same way from Criteria 1 to 6, each expert identified the part that is considered to be a problem in the design or implementation status. After combining the content problems identified by the five experts, the same experts evaluate the severity and priority of the content problems.

- User evaluation of CIMSCITY 2.0: First, a total of five users who were familiar with building LCA were recruited online or through a bulletin board for user evaluation using the thinking aloud test. Second, a total of five evaluators who were members of Korea Internet Professional Association (KIPFA), a non-profit organization consisting of Internet-related experts, were recruited for user evaluation using heuristic evaluation [62]. Both the five users and five evaluators evaluated the usability of CIMSCITY 2.0 in accordance with the user evaluation procedure. Starting from Login page of CIMSCITY 2.0, both the five users and five evaluators identify their own usability related problems discovered while using each page of CIMSCITY 2.0 (refer to Fig. 5). As each page includes several sub-pages, the user or evaluator indicates where the usability related problem has occurred specifically from the sub-pages and identifies the problem in detail (refer to Table 2). Since the five users are not experts in usability, the usability problems identified by the five users are passed on to the five evaluators. After combining the usability problems identified by both the users and evaluators, evaluators define the usability attribute (refer to Table 3) violated by each usability problem, and evaluate the severity and priority of the corresponding usability problem.

4.1. The content evaluation and interpretation for CIMSCITY 2.0

A total of five experts participated in the evaluation of the contents of CIMSCITY 2.0 by stage, and a total of 48 content status and 46 content problems were derived from the evaluation results. Based on the 48 content status and 46 content problems, the content evaluation results of CIMSCITY 2.0 were interpreted in terms of three aspects, as follows: (i) Content status evaluation and interpretation based on the S-I analysis; (ii) Content problem evaluation and interpretation based on the S-P analysis; and (iii) Content improvement strategy for CIMSCITY 2.0.

4.1.1. Content status evaluation and interpretation based on the S-I analysis

The S-I analysis results for the two types of content status, the design and implementation status, according to the six evaluation criteria are shown in Table 8 and Fig. 6. As a result, 24 design status and 24 implementation status were identified, resulting in a total of 48 content status. For the 24 design status, the mean satisfaction value was 4.10 while the mean importance value was 4.81, which suggests that overall, the satisfaction with the design status of CIMSCITY 2.0 was high, and its level of importance was also very high. In the case of the 24 implementation status, the mean satisfaction value was 3.95 while the mean importance value was 4.84, indicating that both the importance of and satisfaction with the implementation status of CIMSCITY 2.0 were high, but a lower satisfaction value was shown compared to the design status. In both the design and implementation status, the evaluation criterion that had the highest number of content status included in Quadrant II, requiring top-priority improvement, was found to be Criterion 4 while the evaluation criteria with the highest number of content status included in Quadrants I and IV, requiring no improvement, were Criteria 1 and 2.

Table 8. The Satisfaction-Importance (S-I) analysis result of content status

Content Status	Criteria	Number of content status				Total
		Quadrant II	Quadrant III	Quadrant IV	Quadrant I	
		(High[a] →	Low[b],	N/A[c]	)	
Design status	Criterion 1	0	0	4	0	4
	Criterion 2	0	0	1	3	4
	Criterion 3	2	0	0	2	4
	Criterion 4	4	0	0	0	4
	Criterion 5	1	0	0	3	4
	Criterion 6	2	1	1	0	4
	Total	9 (38%)	1 (4%)	6 (25%)	8 (33%)	24 (100%)
Implemen-tation status	Criterion 1	0	0	4	0	4
	Criterion 2	0	0	1	3	4
	Criterion 3	0	1	0	3	4
	Criterion 4	4	0	0	0	4
	Criterion 5	4	0	0	0	4
	Criterion 6	2	0	2	0	4
	Total	10 (42%)	1 (4%)	7 (29%)	6 (25%)	24 (100%)

Note: High[a] represents high improvement priority; Low[b] represents low improvement priority; and N/A[c] represents improvement unnecessary.

First, the content status included in Quadrant II, requiring top-priority improvement, were found to be the ones corresponding to Criteria 4, 5, and 6 related to assessment, result, and reference. In particular, as all the design and implementation status corresponding to Criterion 4, representing the assessment method, were included in Quadrant II, they were considered by the experts' top priorities in terms of improvement.

Second, the content status included in Quadrant III, requiring second-level priority improvement, was found to be present in both the design and implementation status. The analysis also revealed that for both the design status corresponding to Criterion 6, representing the assessment reference, and the implementation status falling under Criterion 3, representing the specific purpose of the tool, improvement is necessary albeit not urgent.

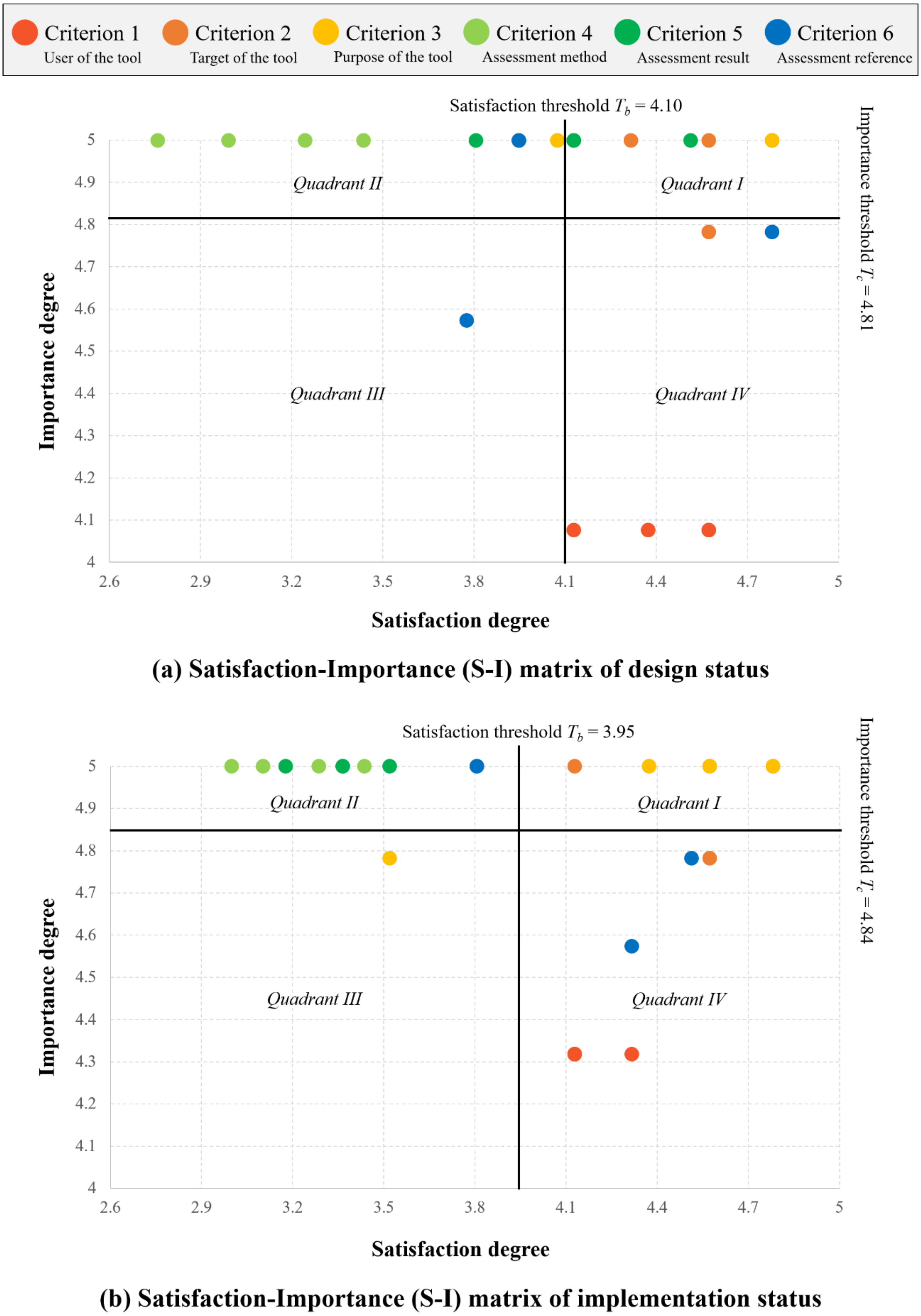


Fig. 6. The Satisfaction-Importance (S-I) matrix of content status by evaluation criteria

Third, the content status included in Quadrants I and IV, requiring no improvement, were the ones corresponding to Criteria 1 and 2 related to the user and target of the tool. All the design and implementation status corresponding to Criterion 1, representing the user of the tool, were found to be included in Quadrant IV. This indicates that the status itself, whose functions are excessively implemented due to their higher satisfaction compared to their importance, is satisfactory albeit less important than the other criteria. Meanwhile, the content status falling under Criteria 2 and 3 were included in Quadrant I, which require no improvement because of their high satisfaction and importance. In particular, as all the design and implementation status corresponding to Criterion 2 were included in Quadrants I and IV, the target of the tool was found to be rationally designed and well implemented.

In summary, as their levels of importance were higher than those of the other contents of CIMSCITY 2.0, the content status that needs to be improved most preferentially were mainly included in Criteria 4, 5, and 6. The content status included in Criteria 2 and 3 were found to have higher satisfaction and importance than those included in the other criteria, indicating that the function is best implemented in CIMSCITY 2.0, and the content status included in Criterion 1 was rated by the experts as the least important item.

4.1.2. Content problem evaluation and interpretation based on the S-P analysis

The S-P analysis results for the two types of content problem, design, and implementation problems, according to the six evaluation criteria are shown in Table 9. As a result, 11 design problems and 35 implementation problems were identified, resulting in a total of 46 content problems. The mean severity value of the 11 design problems was 3.09, and the mean priority value was 3.75, which suggests that there is a need for improvement, but the excessive investment of resources is not required. For the 35 implementation problems, smaller amounts of resources were required for improvement compared to the design problems as the mean severity value was 2.41 and the mean priority value was 3.53. Most of the design and implementation problems were found to be related to Criterion 4. Among the content problems by quadrant, the most dominant ones were also related to Criterion 4, followed by the content problems related to Criterion 5, and design and implementation problems occurred in all the quadrants, as with Criterion 4. Also, only design problems were found in Criterion 6 while only implementation problems were found in Criterion 2. Meanwhile, the least number of content problems occurred in Criteria 1 and 3, and no content problems occurred in Criterion 1 (refer to Table 9 and Fig. 7).

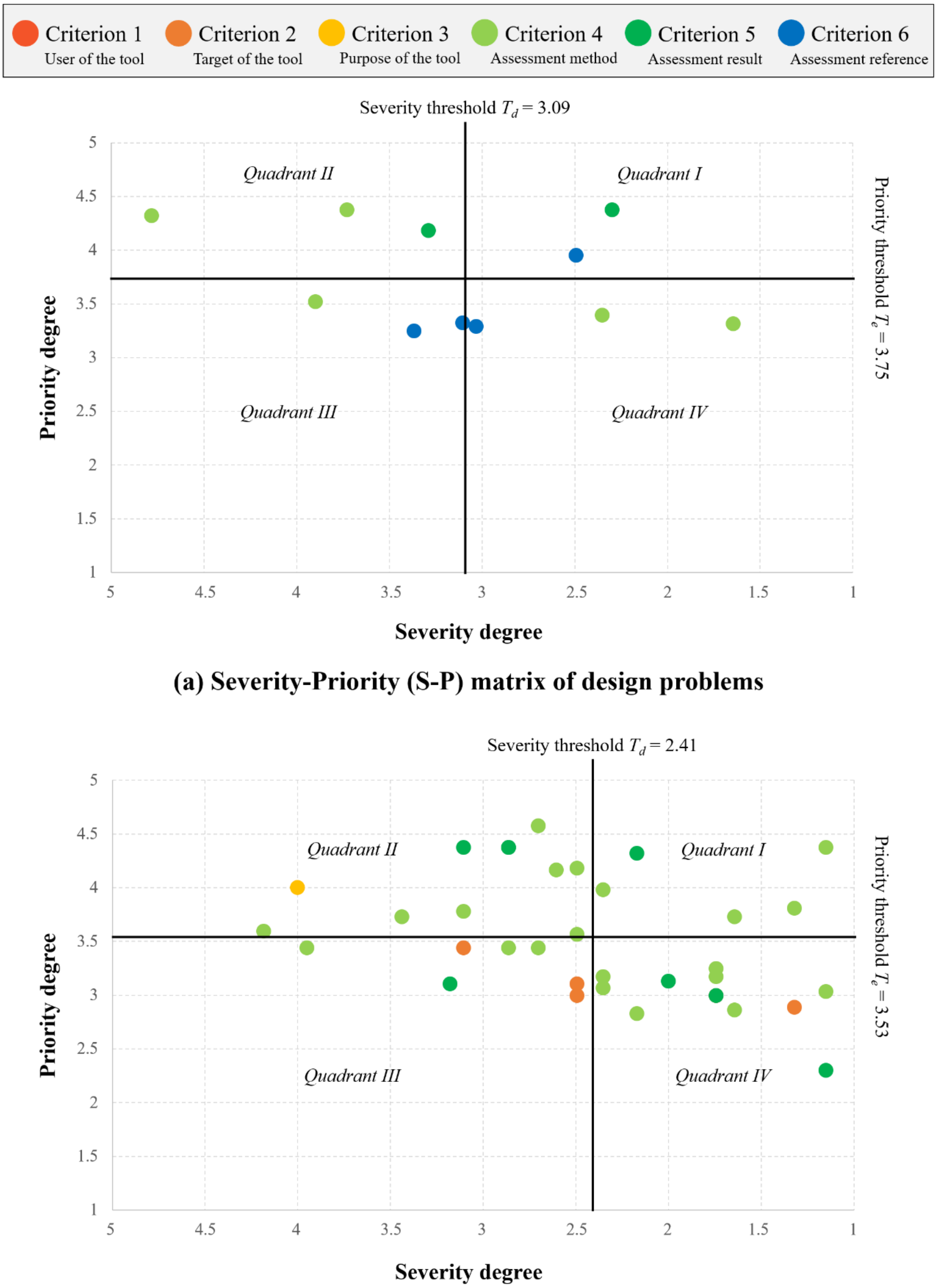
Table 9. The Severity-Priority (S-P) analysis result of content problems

Content Problem	Criteria	Number of content problems				
		Quadrant II (High[a]	Quadrant I	Quadrant IV →	Quadrant III Low[b])	Total
Design Problem	Criterion 1	0	0	0	0	0
	Criterion 2	0	0	0	0	0
	Criterion 3	0	0	0	0	0
	Criterion 4	2	0	2	1	5
	Criterion 5	1	1	0	0	2
	Criterion 6	0	1	1	2	4
	Total	3 (27%)	2 (18%)	3 (27%)	3 (27%)	11 (100%)
Implemen-tation Problem	Criterion 1	0	0	0	0	0
	Criterion 2	0	0	2	3	5
	Criterion 3	1	0	0	0	1
	Criterion 4	7	4	7	3	21
	Criterion 5	3	1	3	1	8
	Criterion 6	0	0	0	0	0
	Total	10 (42%)	1 (4%)	7 (29%)	6 (25%)	24 (100%)

Note: High[a] represents high improvement priority; and Low[b] represents low improvement priority.

The results of the analysis by improvement priority are as follows. First, as the content problems related to Criteria 4 and 5 were mainly included in Quadrant II, with the highest improvement priority, and Quadrant I, with the second-level improvement priority, it was confirmed that the assessment method and results should be intensively improved. In particular, the analysis revealed that in the S-I analysis results for the content status, the content status corresponding to Criterion 4, representing the assessment method, requires the most urgent improvement. As in the S-P analysis results for the content problems, the highest number of content problems corresponding to Criterion 4 were included in Quadrant II, indicating a critical need for and urgency of improvement. As the content problems corresponding to Quadrants II and I should be corrected unconditionally, they need to be considered in the order of higher amounts of resources put into improvement. In this connection, it was found that the most critical problem with the highest level of severity was the implementation problem regarding the assessment method of the construction stage. This problem was that the input data entered in a specific process were unable to save.

Second, the content problems related to Criteria 4 and 5 as well as those corresponding to Criteria 2 and 6 were included in Quadrant IV, requiring no urgent improvement but ensuring easy correction, and Quadrant III, posing difficulties in correction despite the low need for improvement, and this confirmed the need for improvement of the target tool and assessment reference. Meanwhile, the S-I analysis results for the content status showed that almost no improvement is required for the content status related to Criterion 2, but the S-P analysis results for the content problems confirmed a need for improvement, although the improvement is not as urgent as that for the implementation problems related to Criterion 2, corresponding to Quadrant III or IV. This is because the experts expressed that they were satisfied regardless of whether there were problems as they considered Criterion 2 less important than the other criteria. On the contrary, the S-I analysis results for the content status showed that improvement is required for the content status related to Criterion 6, but the S-P analysis results for the content problems confirmed only the design problems related to Criterion 6, requiring no urgent improvement as they correspond to Quadrant III or IV. This is because the experts responded that these should be considered in the improvement process regardless of whether there were problems or none as they considered Criterion 6 more important than Criterion 2. As the need for improvement is low for the content problems corresponding to Quadrants IV and III, these problems should be considered in the order of smaller amounts of resources put into improvement. In this connection, it was found that the problem with a low level of severity was the implementation problem regarding the target tool in the D&D stage. This problem was that an exact expression of the building usage was not reflected in the system.



(a) Severity-Priority (S-P) matrix of design problems

(b) Severity-Priority (S-P) matrix of implementation problems

Fig. 7. The Severity-Priority (S-P) matrix of content problems by evaluation criteria

In summary, the top-priority content problems for improvement due to their high priority in CIMSCITY 2.0 were mainly included in Criteria 4 and 5. In Criteria 2 and 6, the content problems can be considered second-priority problems because the content problems themselves occur less frequently, or improvement is not urgent. It was confirmed, however, that the experts considered Criterion 6 relatively more important, and Criterion 2 relatively less important. Lastly, one content problem related to Criterion 3 occurred while no content problem related to Criterion 1 occurred.

4.1.3. Content improvement strategy for CIMSCITY 2.0

Overall, the evaluation revealed that in CIMSCITY 2.0, improvement is required mainly for the content status and problems related to Criteria 4 and 5 and the assessment method and result, and there were more design problems than implementation problems. As the improvement direction and method for CIMSCITY 2.0 composed of four stages according to the building life cycle may vary depending on the stage, specific content improvement stages were determined considering both the evaluation criteria and stage of the tool (refer to Fig. 8).

First, as shown in Fig. 8 (a), the stages with the highest number of content problems to be solved more preferentially among the four stages of the tool are the construction and D&D stages. In the case of the construction stage, the largest amount of resources should be invested so that the assessment method would be well implemented, while in the case of the D&D stage, a large amount of resources should be put into improvement so that the assessment result could be well displayed. One of the possible improvement strategies in the construction stage is directly linking the data input method to the building information modeling (BIM) based system to allow data to be entered automatically rather than entering data manually, while among the possible improvement strategies in the D&D stage are changing the configuration of the result window and improving the graph to display the results in various ways.

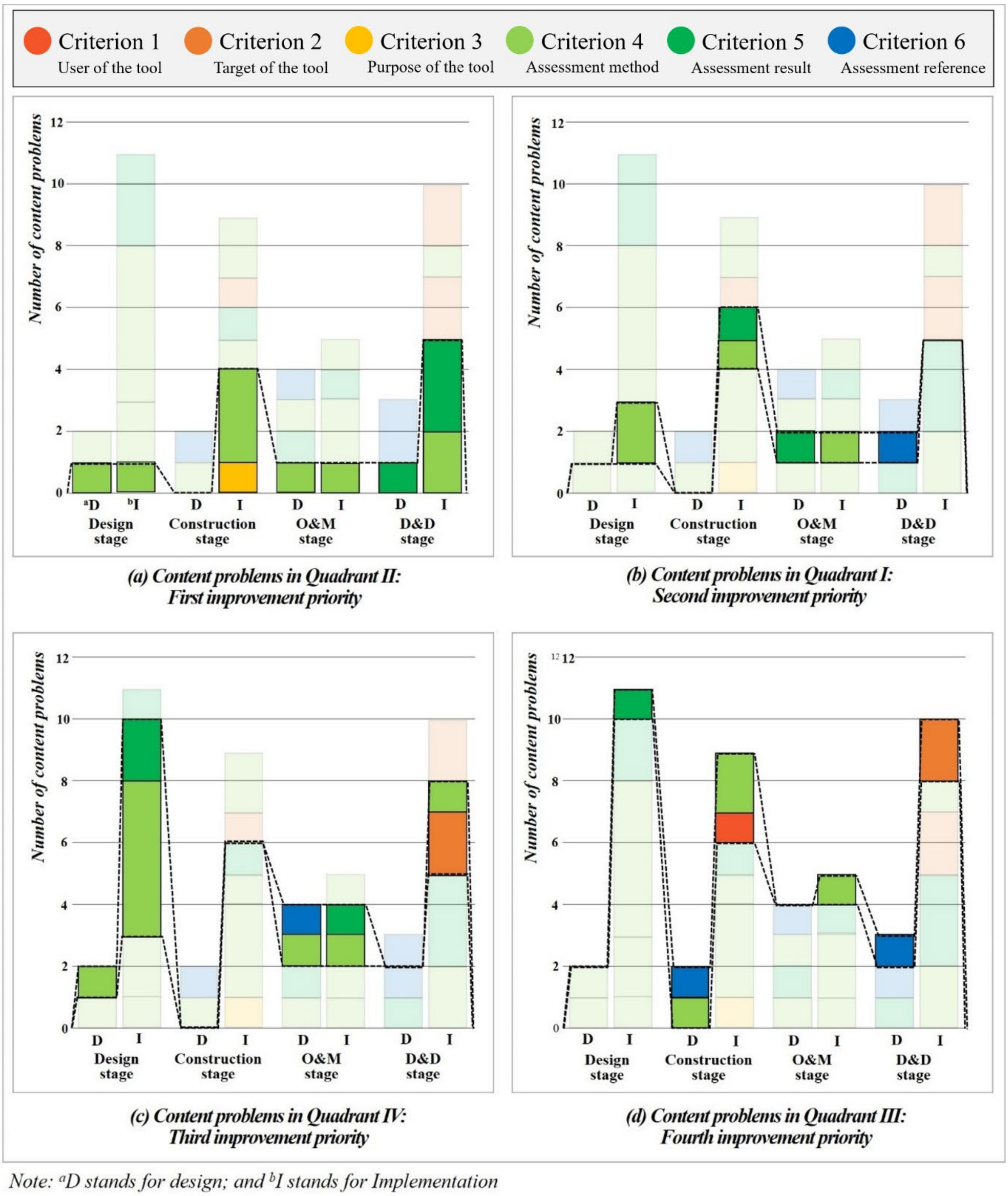


Fig. 8. Content problems by evaluation criteria, improvement priority, and stage of the tool

Second, as shown in Fig. 8 (c), the highest number of content problems occurred in the design stage among the four stages of the tool. Most of the content problems, however, could be improved with a small amount of resources because their severity levels were low. One of the possible improvement strategies in the design stage is additionally writing the unit representation of variables and result values.

Third, as shown in Fig. 8, the smallest number of content problems occurred in the O&M stage among the four stages of the tool, and more design problems occurred in the other stages, requiring improvement in both the design and implementation problems. One of the possible improvement strategies in the O&M stage is adding a new result window that can display the LCA and LCC results at the same time.

4.2. The usability evaluation and interpretation for CIMSCITY 2.0

A total of five users and five evaluators participated in the evaluation of the four stages of CIMSCITY 2.0, and a total of 126 usability problems were derived from the evaluation results. Based on the 126 usability problems, the usability evaluation results for CIMSCITY 2.0 were interpreted in terms of three aspects, as follows: (i) usability problem evaluation and interpretation based on the usability attributes; (ii) usability problem evaluation and interpretation based on S-P analysis; and (iii) usability improvement strategy for CIMSCITY 2.0.

4.2.1. Usability problem evaluation and interpretation based on the usability attributes

The overall usability problem evaluation results for the six UI pages of the target tool according to the six usability attributes are shown in Table 10. As shown in Table 10, usability problems rarely occurred in the UI pages without an analysis function, as in the login or main pages, while they occurred evenly throughout the four stages in the remaining UI pages of the four stages and in the design and D&D stages, which require various analysis functions. First, the usability attribute with the highest number of usability problems was found to be satisfaction, with its usability problems accounting for 41% of the total. This is because CIMSCITY 2.0 is a building LCA tool developed for experts, and thus, the satisfaction felt by general users when using the target tool is not sufficiently taken into account. In particular, the evaluators pointed out that the images on the main UI pages were awkwardly placed, and thus, most of the problems related to the main UI pages were found to violate user satisfaction.

Table 10. The usability problem evaluation result by usability attributes

Usability Attribute	Number of usability problems						
	Login	Main	Design stage	Construction stage	O&M stage	D&D stage	Total
Effectiveness	0	0	1	2	2	1	6 (5%)
Efficiency	0	0	5	4	8	3	20 (16%)
Satisfaction	1	3	12	9	14	13	52 (41%)
Learnability	0	1	10	15	4	4	34 (27%)
Memorability	1	0	1	1	0	1	4 (3%)
Errors	0	0	2	1	3	4	10 (8%)
Total	2 (2%)	4 (3%)	31 (25%)	32 (25%)	31 (25%)	26 (21%)	126 (100%)

Second, the usability attribute with the second-highest number of usability problems was learnability, with its usability problems accounting for 27% of the total. This is because CIMSCITY 2.0, which was developed for experts, does not consider the general users in the understanding and learning processes for identifying the data input and analysis procedures and results. In particular, the evaluators pointed out that the English abbreviations on the UI page of the construction stage, which are difficult for the users to understand, are among the usability problems related to learnability.

Third, the usability attribute with the third-highest number of usability problems was efficiency, with its usability problems accounting for 16% of the total. This was due to the insufficient consideration of the complexity of the input data due to the characteristics of the building LCA tool. In particular, 40% of the usability problems corresponding to efficiency occurred in the UI page of the O&M stage, in which enormous data input qualities are present. Also, the evaluators discovered the problem that when a search item is entered in the UI page of the O&M stage, a search is done, but the items that are being searched are not known.

Lastly, the usability problems related to errors, effectiveness, and memorability accounted for less than 10% of the total. As CIMSCITY 2.0 was developed considering its reliability and repetitive use by building LCA experts, there was almost no problem in obtaining the desired building LCA results from the use of the tool. Also, errors rarely occurred, or only errors that could be dealt with occurred. The analysis also revealed that the tool is implanted to make it easy for the user to remember the procedures, operation methods, and functions when using the tool again, to ensure the repetitive use of the tool.

4.2.2. Usability problem evaluation and interpretation based on the S-P analysis

The S-P analysis results for the 126 usability problems according to the six usability attributes are shown in Table 11. The mean severity value of the 126 usability problems was 2.64, while the mean priority value was 2.95, which suggests that there is a need for overall improvement, but the excessive investment of resources is not required. The usability attribute with the highest number of usability problems included in Quadrant II, requiring top-priority improvement, was found to be learnability, while the usability attribute with the highest number of usability problems but is included in Quadrants III and IV, with low-priority improvement, was found to be satisfaction, when most of the usability problems (83%) are included in Quadrant II, with the highest improvement priority, in which both the severity and priority are high, or in Quadrant IV, where improvement can be achieved with a small amount of resources due to the low severity and priority (refer to Table 11 and Fig. 9).

Table 11. The Severity-Priority (S-P) analysis result of usability problems

Usability Attributes	Number of usability problems				
	Quadrant II (High[a]	Quadrant I	Quadrant IV →	Quadrant III Low[b])	Total
Effectiveness	4	0	2	0	6 (5%)
Efficiency	11	2	6	1	20 (16%)
Satisfaction	4	4	39	5	52 (41%)
Learnability	21	4	5	4	34 (27%)
Memorability	2	0	1	1	4 (3%)
Errors	8	1	1	0	10 (8%)
Total	50 (40%)	11 (9%)	54 (43%)	11 (9%)	126 (100%)

Note: High[a] represents high improvement priority; and Low[b] represents low improvement priority.

First, the usability problems corresponding to learnability, efficiency, and errors were mainly included in Quadrant II, requiring the highest improvement priority. Of these three usability attributes, the usability problems corresponding to learnability occurred most frequently in Quadrant II. It was found that although some of the usability problems included in Quadrant II showed relatively low severity and priority (the light-green circles near the bottom right corner of Quadrant II in Fig. 9), the largest amounts of resources were still required for improvement. Meanwhile, one of the most critical usability problems that violate learnability occurred in the UI page of the construction stage, and the evaluators pointed out the problem that the function of adding or deleting data could not be recognized. After learnability, efficiency was found to have the highest number of usability problems occurring in Quadrant II, and showed relatively low severity and priority among the usability problems included in Quadrant II (the orange circles near the bottom right corner of Quadrant II in Fig. 9). Also, one of the most critical usability problems that violate efficiency occurred in the UI page of the construction stage, and the evaluators pointed out that the arrow keys for moving the cursor up and down to see the top and bottom of the table were inefficiently placed. Lastly, among the three usability attributes, errors exhibited the smallest number of usability problems occurring in Quadrant II, but showed relatively high severity and priority among the usability problems included in Quadrant II because it is impossible to use the building LCA tool when the analysis function does not work due to the occurrence of errors (the blue circles near the upper left corner of Quadrant II in Fig. 9). Also, one of the most critical usability problems that violate errors occurred in the UI page of the D&D stage, and the evaluators discovered the problem that data are displayed even when data cannot be loaded onto the tool and there are no loaded data.

Second, the usability problems corresponding to satisfaction and learnability were mainly included in Quadrant I, with the second-highest improvement priority. Satisfaction and learnability were the top two usability attributes where usability problems occurred most frequently, and the focus should be placed on the efficient improvement of the said usability problems. As the usability problems included in Quadrant I have high priority but low severity, improvement can be made even with a small amount of resources. Therefore, the improvement of these usability problems makes it possible to effectively increase the satisfaction and learnability.

Third, the usability problems corresponding to satisfaction, efficiency, and learnability were mainly included in Quadrant IV, requiring no urgent improvement but ensuring simple correction. In particular, satisfaction was the usability attribute that had the largest number of usability problems, but because most of the problems were included in Quadrant IV, many of them were expected to be improved more easily. As one of the usability problems with low severity that violate satisfaction occurred in the UI page of the O&M stage, a problem arises with unnecessary text boundaries on the information displayed on the page.

Lastly, the usability problems corresponding to satisfaction and learnability were mainly included in Quadrant III, posing difficulties in correction despite the low need for improvement. As mentioned above, it can be confirmed that as satisfaction and learnability were the top two usability attributes with the largest number of usability problems, the said usability problems were also included in Quadrant III. As the usability problems included in Quadrant III, however, had low priority but showed high severity, the efficiency of improvement significantly decreased. Therefore, to increase the satisfaction and learnability, it is effective to focus on the usability problems included in the other quadrants rather than on those included in the above.

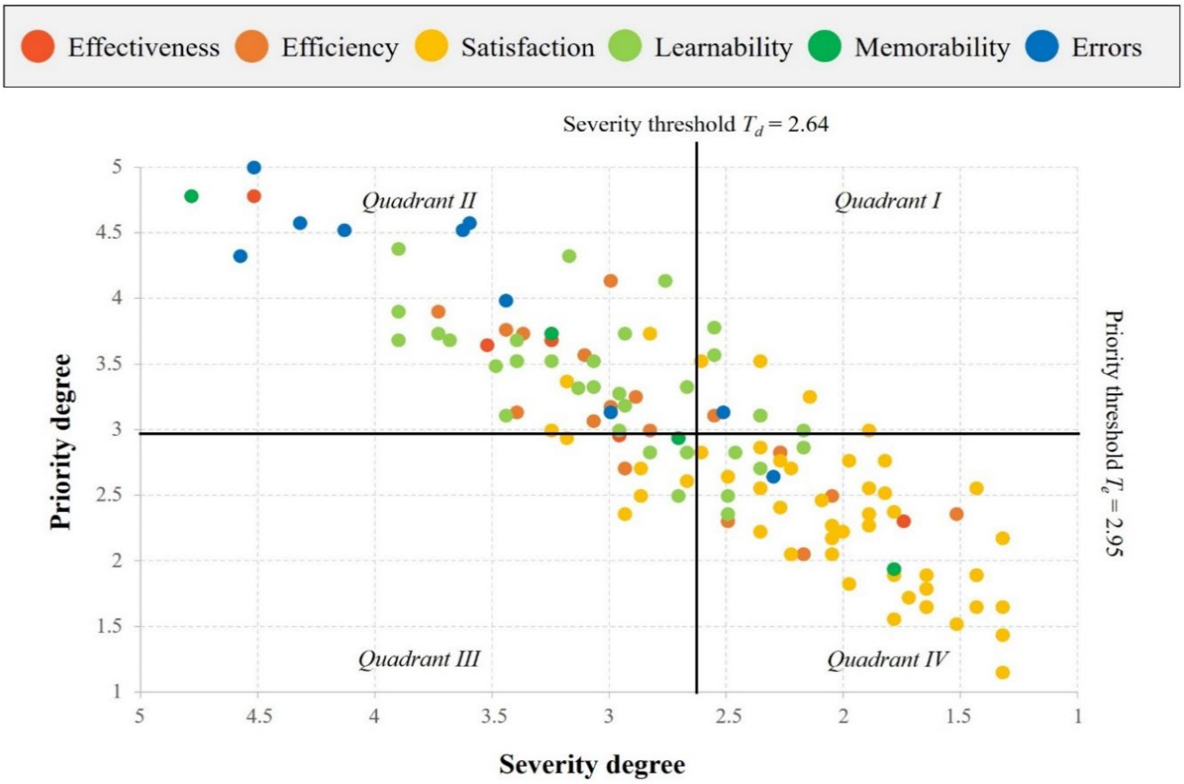


Fig. 9. The Severity-Priority (S-P) matrix of usability problems by usability attributes

In summary, the usability problems that need to be improved most preferentially as their priority and severity are high in CIMSCITY 2.0 were found to be the ones corresponding to learnability, efficiency, and errors. In satisfaction, the largest number of usability problems occurred, but the analysis revealed that they do not require urgent improvement and ensure easy correction. The usability attributes with the smallest number of usability problems and the lowest need for improvement were found to include effectiveness and memorability.

4.2.3. Usability improvement strategy for CIMSCITY 2.0

Overall, a large number of usability problems corresponding to learnability and satisfaction occurred in CIMSCITY 2.0, which has no problem in achieving the desired result (i.e., effectiveness) and using it after a long time (i.e., memorability), and leads to almost no problem (errors) but rarely considers use by non-experts. Based on these analysis results, specific usability improvement strategies were determined according to the usability attributes.

First, usability problems themselves occurred frequently in the order of satisfaction, learnability, efficiency, errors, effectiveness, and memorability, among which the usability problem corresponding to learnability requires the largest amount of resources for improvement. Also, it was found that various critical usability problems occurred in the learning process for the new users' use of CIMSCITY 2.0 because CIMSCITY 2.0 lacks consideration of its new users. Meanwhile, one of the possible improvement strategies for learnability is reorganizing the order of proceeding with the building LCA in CIMSCITY 2.0 to facilitate the users' understanding.

Second, a large amount of resources should be invested to improve the usability problems corresponding to efficiency and error, after learnability. In CIMSCITY 2.0, there are many data inputs, and the procedures are complicated. Therefore, one of the possible improvement strategies for efficiency is enlarging the input window so that the input data can be checked at once even if its name is long. In the case of errors, usability problems did not occur frequently but showed the highest severity and priority, which indicates that the need for and urgency of improvement is critical. If an error occurs in the process of an important analysis of the building LCA tool and the analysis function does not work properly, the reliability of the use of the building LCA tool decreases. Therefore, one of the possible improvement strategies for errors is planning the provision of a guide window to inform the users of the specific procedures for solving the errors.

Third, the largest number of usability problems occurred in satisfaction, but their severity was low, which suggests that most of the usability problems can be improved even with a small amount of resources. In this regard, two possible improvement strategies for satisfaction are changing the font size and graph color and adding icons.

5. Conclusions

Until recently, there have been many studies on the assessment method of building life cycle assessment (LCA) tools for analyzing and diagnosing buildings and for improving buildings' retrofit problems. No research has been conducted, however, on the evaluation process for analyzing and diagnosing the LCA tool itself and retrofit problems. In this regard, this study proposed an evaluation framework that could be applied to a single building LCA tool.

Through a case study, the critical problems in the contents and usability of the target building LCA tool (i.e., Carbon-integrated Construction Management System in CITY [CIMSCITY 2.0]) were derived and analyzed, and the results confirmed that the evaluation framework proposed in this study could actually be applicable. If the evaluation results are interpreted using the satisfaction-importance (S-I) and severity-priority (S-P) matrices, the improvement priorities for the contents and usability of the building LCA tool can be determined to achieve effective improvement. Also, if the evaluation is performed considering both the contents and usability of the tool, it is possible not only to prevent improvements made in a biased direction towards experts or non-experts from occurring but also to keep other critical problems from occurring in the improvement process. In the development of a professional tool, the maintenance and sustainability of the tool are of great importance. If the current status of the building LCA tool is diagnosed and its shortcomings are conquered through this systematic evaluation and improvement method, the users' continued use of the tool can be ensured by increasing the tool's reliability and usability.

This study was qualitative research focused on discovering the problems of the contents and usability of the target building LCA tool, and on its evaluation and improvement. For the future work, quantitative research that quantitatively compares the analysis performance (e.g., usage time, frequency of use, number of errors) before and after the improvement of a certain building LCA tool needs to be conducted to verify how useful the evaluation framework presented in the study is. Moreover, it is expected that if many existing building LCA tools are analyzed and diagnosed using the evaluation framework proposed in this study, it would be possible to generalize key factors for developing and improving different building LCA tools. The generalized key factors would be useful and can be referenced when organizations such as ISO or World Green Building Council decide to establish international, national, or localized standards and procedures for evaluating and certifying building LCA tools. Comparative evaluation between different building LCA tools can then be enabled based on the established standards and procedures.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT; Ministry of Science and ICT) (NRF-2018R1A2A1A19020868).

References

[1] UN Environment and International Energy Agency. Towards a zero-emission, efficient, and resilient buildings and construction sector. Global Status Report 2019. 2019.

[2] Olubunmi OA, Xia PB, Skitmore M. Green building incentives: A review. Renew Sustain Energy Rev 2016. https://doi.org/10.1016/j.rser.2016.01.028.

[3] Zuo J, Zhao ZY. Green building research-current status and future agenda: A review. Renew Sustain Energy Rev 2014. https://doi.org/10.1016/j.rser.2013.10.021.

[4] Wang T, Seo S, Liao PC, Fang D. GHG emission reduction performance of state-of-the-art green buildings: Review of two case studies. Renew Sustain Energy Rev 2016. https://doi.org/10.1016/j.rser.2015.11.037.

[5] Anand CK, Amor B. Recent developments, future challenges and new research directions in LCA of buildings: A critical review. Renew Sustain Energy Rev 2017. https://doi.org/10.1016/j.rser.2016.09.058.

[6] Cabeza LF, Rincón L, Vilariño V, Pérez G, Castell A. Life cycle assessment (LCA) and life cycle energy analysis (LCEA) of buildings and the building sector: A review. Renew Sustain Energy Rev 2014. https://doi.org/10.1016/j.rser.2013.08.037.

[7] Ji C, Hong T, Jeong K, Leigh SB. A model for evaluating the environmental benefits of elementary school facilities. J Environ Manage 2014. https://doi.org/10.1016/j.jenvman.2013.11.022.

[8] Forsberg A, von Malmborg F. Tools for environmental assessment of the built environment. Build Environ 2004. https://doi.org/10.1016/j.buildenv.2003.09.004.

[9] Kim J, Koo C, Kim CJ, Hong T, Park HS. Integrated CO2, cost, and schedule management system for building construction projects using the earned value management theory. J Clean Prod 2015. https://doi.org/10.1016/j.jclepro.2014.05.031.

[10] Baek C, Park SH, Suzuki M, Lee SH. Life cycle carbon dioxide assessment tool for buildings in the schematic design phase. Energy Build 2013. https://doi.org/10.1016/j.enbuild.2013.01.025.

[11] Singh A, Berghorn G, Joshi S, Syal M. Review of life-cycle assessment applications in building construction. J Archit Eng 2011. https://doi.org/10.1061/(ASCE)AE.1943-5568.0000026.

[12] BRE group. IMPACT 2020.

[13] KT Innovations. tally n.d. https://choosetally.com/ (accessed March 21, 2020).

[14] Thinkstep. GaBi software, http://www.gabi-software.com; PE Int 2016.

[15] KieranTimberlake. Tally n.d.

[16] Athena Sustainable Materials Institute. Impact Estimator for Buildings n.d.

[17] Al-Ghamdi SG, Bilec MM. Green Building Rating Systems and Whole-Building Life Cycle Assessment: Comparative Study of the Existing Assessment Tools. J Archit Eng 2017;23:04016015. https://doi.org/10.1061/(asce)ae.1943-5568.0000222.

[18] simapro manual PRe Consultants. Introduction to LCA with SimaPro 7. PRé Consult Netherlands Version 2008.

[19] Shimizu Construction. Global Environmental Model/Management-21P (GEM- 21P) n.d.

[20] Kumanayake R, Luo H. A tool for assessing life cycle CO2 emissions of buildings in Sri Lanka. Build Environ 2018. https://doi.org/10.1016/j.buildenv.2017.11.042.

[21] Haapio A, Viitaniemi P. A critical review of building environmental assessment tools. Environ Impact Assess Rev 2008. https://doi.org/10.1016/j.eiar.2008.01.002.

[22] Gagnon MP, Ngangue P, Payne-Gagnon J, Desmartis M. M-Health adoption by healthcare professionals: A systematic review. J Am Med Informatics Assoc 2016. https://doi.org/10.1093/jamia/ocv052.

[23] Buchanan S, Salako A. Evaluating the usability and usefulness of a digital library. Libr Rev 2009. https://doi.org/10.1108/00242530910997928.

[24] Aranda-Jan CB, Mohutsiwa-Dibe N, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. BMC Public Health 2014. https://doi.org/10.1186/1471-2458-14-188.

[25] Kang H, Park H-A. A Mobile App for Hypertension Management Based on Clinical Practice Guidelines: Development and Deployment. JMIR MHealth UHealth 2016. https://doi.org/10.2196/mhealth.4966.

[26] Gerber BS, Brodsky IG, Lawless KA, Smolin LI, Arozullah AM, Smith E V., et al. Implementation and evaluation of a low-literacy diabetes education computer

multimedia application. Diabetes Care 2005. https://doi.org/10.2337/diacare.28.7.1574.

[27] Boberg A, Monis-Khoo S. The Delphi method: A review of methodology and an application in the evaluation of a higher education program. Can J Progr Eval 1992.

[28] Singh MD. Evaluation framework for nursing education programs: Application of the CIPP model. Int J Nurs Educ Scholarsh 2004. https://doi.org/10.2202/1548-923X.1023.

[29] Milward J, Khadjesari Z, Fincham-Campbell S, Deluca P, Watson R, Drummond C. User Preferences for Content, Features, and Style for an App to Reduce Harmful Drinking in Young Adults: Analysis of User Feedback in App Stores and Focus Group Interviews. JMIR MHealth UHealth 2016. https://doi.org/10.2196/mhealth.5242.

[30] Kim CJ, Kim J, Hong T, Koo C, Jeong K, Park HS. A program-level management system for the life cycle environmental and economic assessment of complex building projects. Environ Impact Assess Rev 2015. https://doi.org/10.1016/j.eiar.2015.04.005.

[31] Suh S, Huppes G. Missing inventory estimation tool using extended input-output analysis. Int J Life Cycle Assess 2002. https://doi.org/10.1007/BF02994047.

[32] Olsina L, Rossi G. Measuring Web application quality with WebQEM. IEEE Multimed 2002. https://doi.org/10.1109/MMUL.2002.1041945.

[33] Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps. JMIR MHealth UHealth 2015. https://doi.org/10.2196/mhealth.3422.

[34] Garnett C, Crane D, West R, Brown J, Michie S. Identification of Behavior Change Techniques and Engagement Strategies to Design a Smartphone App to Reduce Alcohol Consumption Using a Formal Consensus Method. JMIR MHealth UHealth 2015. https://doi.org/10.2196/mhealth.3895.

[35] Khajouei R, Zahiri Esfahani M, Jahani Y. Comparison of heuristic and cognitive walkthrough usability evaluation methods for evaluating health information systems. J Am Med Informatics Assoc 2017. https://doi.org/10.1093/jamia/ocw100.

[36] Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016. https://doi.org/10.1016/j.jbi.2016.02.002.

[37] Attwood S, Parke H, Larsen J, Morton KL. Using a mobile health application to reduce alcohol consumption: a mixed-methods evaluation of the drinkaware track & calculate units application. BMC Public Health 2017. https://doi.org/10.1186/s12889-017-4358-9.

[38] Lew P, Olsina L, Zhang L. Quality, quality in use, actual usability and user experience as key drivers for web application evaluation. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 2010. https://doi.org/10.1007/978-3-642-13911-6_15.

[39] International Organization for Standardization 14044. Environmental Management. Life Cycle Assessment. Requirements and Guidelines. Ntc-Iso 14044 2007.

[40] Peng C. Calculation of a building's life cycle carbon emissions based on Ecotect and building information modeling. J Clean Prod 2016. https://doi.org/10.1016/j.jclepro.2015.08.078.

[41] Dworkin SL. Sample size policy for qualitative studies using in-depth interviews. Arch Sex Behav 2012. https://doi.org/10.1007/s10508-012-0016-6.

[42] Fernandez A, Insfran E, Abrahão S. Usability evaluation methods for the web: A systematic mapping study. Inf. Softw. Technol., 2011. https://doi.org/10.1016/j.infsof.2011.02.007.

[43] Ivory MY, Hearst MA. The state of the art in automating usability evaluation of user interfaces. ACM Comput Surv 2001. https://doi.org/10.1145/503112.503114.

[44] Van Waes L. Thinking aloud as a method for testing the usability of websites: The influence of task variation on the evaluation of hypertext. IEEE Trans Prof Commun 2000. https://doi.org/10.1109/47.867944.

[45] Krahmer E, Ummelen N. Thinking about thinking aloud: A comparison of two verbal protocols for usability testing. IEEE Trans Prof Commun 2004. https://doi.org/10.1109/TPC.2004.828205.

[46] Nielsen J. Estimating the number of subjects needed for a thinking aloud test. Int J Hum - Comput Stud 1994. https://doi.org/10.1006/ijhc.1994.1065.

[47] Nielsen J. How Many Test Users in a Usability Study? Nielsen Norman Gr 2012.

[48] Nielsen J. How to Conduct a Heuristic Evaluation Nielsen. Nielsen Norman Gr 1995.

[49] Nielsen J, Landauer TK. Mathematical model of the finding of usability problems. Conf. Hum. Factors Comput. Syst. - Proc., 1993. https://doi.org/10.1145/169059.169166.

[50] Jaspers MWM. A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence. Int J Med Inform 2009. https://doi.org/10.1016/j.ijmedinf.2008.10.002.

[51] Nielsen J. 10 Usability Heuristics for User Interface Design. Conf Companion Hum Factors Comput Syst CHI 94 1995. https://doi.org/10.1145/191666.191729.

[52] Nielsen J. Usability Engineering. Boston: Acad Press 1993.

[53] Nielsen J. Designing Web Usability New Riders Publishing. Indianap USA 2000.

[54] Standard I. ISO 9241-210. Int Organ 2009. https://doi.org/10.1021/es0620181.

[55] Geng X, Chu X. A new importance-performance analysis approach for customer satisfaction evaluation supporting PSS design. Expert Syst Appl 2012. https://doi.org/10.1016/j.eswa.2011.08.038.

[56] Guadagnolo F. The importance-performance analysis: An evaluation and marketing tool. J Park Recreat Admi 1985;3(2).

[57] Martilla JA, James JC. Importance-Performance Analysis. J Mark 1977. https://doi.org/10.2307/1250495.

[58] Yang LJ, Chou TC, Ding JF. Using the Importance-Performance Analysis (IPA) approach to measure the service quality of mobile application stores in Taiwan. African J Bus Manag 2011. https://doi.org/10.5897/AJBM10.1163.

[59] Sever I. Importance-performance analysis: A valid management tool? Tour Manag 2015;48:43–53. https://doi.org/10.1016/j.tourman.2014.10.022.

[60] Zhang H, Gong L, Versteeg S. Predicting bug-fixing time: An empirical study of commercial software projects. Proc. - Int. Conf. Softw. Eng., 2013. https://doi.org/10.1109/ICSE.2013.6606654.

[61] Koo C, Hong T, Hyun C, Koo K. A CBR-based hybrid model for predicting a construction duration and cost based on project characteristics in multi-family housing projects. Can J Civ Eng 2010. https://doi.org/10.1139/L10-007.

[62] KIFPA 2020. https://www.kifpa.or.kr.