Usability Information Management: Prototype for Result Exploration Based on an Empirical Analysis of Use Cases

Ben Heuwing, Thomas Mandl, Christa Womser-Hacker

Dept. of Information Science and Natural Language Processing, University of Hildesheim

Hildesheim, Germany

heuwing@uni-hildesheim.de

ABSTRACT

Empirical usability results from user research and usability evaluations are valuable as an information resource beyond the scope of the project for which they were originally collected [22]. This study presents the results of interviews with usability practitioners ($n = 8$) to describe existing use cases for this type of information. Based on the use cases, an information structure for usability information has been created. Its application was discussed in focus groups in two organizations. As a result, an information management system for usability information was implemented and evaluated with users in a qualitative study ($n = 11$). The results indicate the diversity of existing use cases. These include direct application to current design questions, exploration of results with the aim of learning and devising generalized internal guidelines, and analyzing results to support tracking of implementation, controlling, and reporting. Aspects of the product and the respective context of use appear to be essential to support exploration of usability results. In addition, information that helps to make judgments about their reliability is necessary to enable systematic reuse.

Author Keywords

Usability evaluation methods; usability information management; usability information system.

ACM Classification Keywords

H.5.2 [Information Interfaces and Presentation]: *User Interfaces* – Evaluation/methodology; User-centered design

INTRODUCTION

During the process of user-centered design, large software development organizations continuously produce valuable results and insights about the processes and structures in the domain of the product that is being developed as well as knowledge on how to represent those processes and structures

at the level of the interface [9]. As a consequence, the usability-related information [15] that is typically available in an organization includes rich resources describing the context of use, user needs, and requirements as well as reports on evaluation studies and user feedback collected during use. While these results are primarily produced to communicate results during the design process within a single project, they can be applied during the design of new versions and similar products. This has enormous potential to focus user-centered design on the most important questions of a project, can help to base discussions on knowledge available to all stakeholders instead of the view of experts, and will help to avoid the recurrence of usability problems across versions and products. Rosenbaum [22] therefore argues for the implementation of “*knowledge management to improve return on usability investments*” and recommends structuring collected results using adequate metadata to enable usability practitioners to apply them to their current design problems.

Currently, no dedicated support for this process is available in practice. In the context of the User Action Framework (UAF, [1]) and for results of scenario-based design [9], requests have been put forward to create integrated usability repositories supporting the collection and reuse of evaluation results. In addition, the necessity of usability knowledge management is demonstrated by the need for interface patterns and usability guidelines in many organizations. However, a considerable investment of effort is needed at the time of creation of new guidelines. In addition, most collections of interface patterns and usability guidelines are not based on empirical findings and should therefore be complemented by methods that help systematize and use empirical results from user research and usability evaluations.

Consequently, we propose the concept of *usability information management* (UIM) to describe the systematic transfer of various types of results of the user-centered design process from previous projects or studies to current questions and design problems. In the context of this study, we assume that usability information can only be accessed in a meaningful way when indexed with metadata from user studies. As a foundation, this study therefore systematically analyzes and compares current practices in managing usability information across organizations in Germany. Based on the findings, proposals for a faceted classification of usability information were developed and were discussed in

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored.

Copyright is held by the owner/author(s).
NordiCHI '16, October 23-27, 2016, Gothenburg, Sweden
ACM 978-1-4503-4763-1/16/10.
<http://dx.doi.org/10.1145/2971485.2971530>

This work is licensed under a Creative Commons Attribution-ShareAlike International 4.0 License.

detail in focus groups in two different organizations. Finally, a prototype that provides access to a sample collection of usability information was created and was evaluated with usability practitioners.

Our results demonstrate that both ad-hoc solutions as well as systematic use of existing information systems are already employed for usability information management. In addition to metadata about the product, the interface, and the scenario of use, it is important to document aspects of the practical reliability of results. The key result of interest for practitioners consists of the diversity of possible use cases for usability information and the need to take these into account when creating infrastructures for usability information management.

RELATED WORK

General theories about the reuse of design artifacts and knowledge in the research area of Human-Computer Interaction (HCI) include the domain theory [25] and the task-artifact cycle in the context of scenario-based design [3]. More specifically, several approaches classify results of usability evaluations. In the following sections, these are presented and contrasted with examples of the reuse of usability results from empirical studies.

Result Classifications and Tools to Support Reuse

Research in this area has led to the development of several frameworks that can be used to systematically classify usability results. Recent and empirically validated approaches include the *Classification of Usability Problems* (CUP) and the *User Action Framework* (UAF). CUP provides a classification that adds, among other information, details about the stage of the development process at which the issue is solved, the context within the product, a simple classification of the cause of the problem, and an estimate of the potential cost of correcting the problem [27]. These attributes are intended to help software developers prioritize usability issues. The results of a case study suggest that the vocabulary used for classification should be adapted to the context of the particular organization.

The UAF provides a hierarchical classification of usability problems based on their cause [1]. The framework is intended to support the integration of tools for usability engineering and was constructed based on the theory of the interaction cycle. The UAF has been used to improve the diagnosis of usability problems [12]. Classification with the UAF has been found to be sufficiently reliable [1]. However, the time needed to employ the UAF for classification has been found to be comparatively high [11]. *Vizability* is a tool that supports exploration and analysis of a set of usability problems that have been classified with the UAF [17]. The tool employs an interactive visualization of the hierarchical structure of the UAF. A synthesis of these and other approaches [8] has identified three perspectives that can be used as a conceptual framework to develop a classification of usability problems: “*context of use*”, “*design knowledge*”, and “*design activities*”. The author recom-

mends selecting existing models that are suitable for the intended environment to instantiate each of these perspectives.

The frameworks presented try to solve different problems: While the UAF aims to support diagnosis of usability problems, CUPs is meant to improve communication with developers. In addition, UAF and other approaches have been used to compare usability evaluation methods based on categorized results [28]. An approach that explicitly tries to support reuse across projects and product domains for the practical work of an organization is the *Usability Knowledge Management System* [13]. Product-specific user goals (usage scenarios) and generic sub-goals (representing steps of interaction) help to find examples of usability problems, corresponding empirical data, and solutions. Three different use cases are specified for designers as intended users: (1) using existing information about user tasks and the usability of products to support orientation during the conceptual design phase, (2) the application of findings regarding a specific type of interaction during design, and (3) searching for solutions to known problems after a user test that are similar to those found in the test.

Empirical Studies on Practices of Reuse

While a substantial number of methods and tools can help categorize usability results, few studies have reported on practices of reusing usability results. Some examples of the reuse of empirical usability results can be extracted from studies about the practice of usability engineering. First of all, the individual professional experience of experts is an important resource, for example when data from usability evaluations are analyzed [5]. However, problems may arise when experienced usability practitioners base their argumentation only on personal experience, as this may be difficult to refute if the underlying data are not available [14]. Externalized resources of usability information may emerge when results are communicated within an organization (e.g. reports, lists of problems) but are also built intentionally to make selected expertise available [6]. In addition, a small number of case studies report on the potential to systematically derive internal usability guidelines and interface patterns from empirical results, for example [10]. Comparing and benchmarking the usability of products are additional motivations to collect results in a standardized way, for example, based on standards for reporting evaluation results [19].

RESEARCH QUESTIONS AND METHODS

This work uses a multi-method approach to empirically define and validate use cases of reusing usability information from previous studies. Reuse may involve a transfer between the successive versions of one product or between different products and product domains. Based on the results, we designed and validated tools to help systemize and analyze these types of results. In this context, we defined the following research questions (RQ):

- RQ1: Which use cases of collections of usability results exist in practice?
- RQ2: Which criteria are used to search for relevant results?
- RQ3: How can an information system support the application of collected usability results?
- RQ4: To what degree do internal usability practitioners accept the support of an information system for usability information?

Empirical studies on the use of usability methods “in the wild” treat project stakeholders as users of methods and tools. In this sense, “double ethnography” [21] can be used to design tools for designers [24]. Our potential target group was usability practitioners who are working in-house in development organizations. We expected them to have a higher motivation to reuse usability results collected internally because they were responsible for a more homogeneous set of products. Throughout the studies presented here, we concentrated on in-house practitioners who worked at companies in Germany.

The multi-method approach of this study combined qualitative interviews with focus groups, a scenario-based design process, and an evaluation study based on a prototype. Each study was based on the previous efforts and helped to validate their results. First, interviews with usability practitioners provided insights into current practices of managing and reusing collected usability results (RQ1). These insights were discussed in subsequent focus groups to define requirements regarding an optimized information structure to access usability information (RQ2). Based on the results, an interactive prototype of a usability information system was implemented and evaluated in a study with usability engineers to collect further insights about its application in different organizations (RQ3). The prototype focused on the results of user testing, which may include qualitative findings with proposals for redesign as well as quantitative data from questionnaires. In a cognitive walkthrough, this prototype was presented to usability engineers to assess the acceptance of this type of usability information system (RQ4).

INTERVIEW STUDY ON CURRENT REUSE BEHAVIOUR

Explorative, semi-structured interviews with usability practitioners were conducted with the goal of investigating potential use cases of usability information (RQ1). In addition, the results were analyzed regarding possible barriers to reusing usability information to identify areas for improvement.

Methods and Participants

Eight usability practitioners participated in the interviews. They were recruited through open calls during conferences, with open invitations on mailing lists, and through contacts of other participants. Each participant represented a different organization and was responsible for the usability of products developed in that organization. Participants’ work

involved the evaluation and/or design of interactive systems. Two of the participants were responsible for usability teams in their organizations. In addition, all participants were required to be able to describe a case in which they had reused usability information from previous projects. This was the case for all but one of the candidates who were approached for this study. However, self-recruiting may have hidden a larger number of potential participants who either did not have the experience of using existing usability information or may not have approved of the practice of reuse.

At the beginning of the interviews, the participants were asked to recount a recent project in which they had used existing usability-related information. This helped to ground the discussion in a detailed recount of a specific situation from the everyday work life of the participants. Next, the interviewer proposed other possible use cases for usability information collected from literature and previous interviews and discussed them regarding their relevance in the experience of the participant. At the end of the interview, possible advantages of and barriers to reusing usability information were discussed.

For analysis, all recordings were transcribed. The transcripts were coded according to possible use cases of usability information as well as the perceived advantages and disadvantages of and possible obstacles to its application. A preliminary coding system was established beforehand based on literature research and was continuously revised during the analysis process, thus combining deductive and inductive coding.

Results

The results of this study demonstrated that collected usability results were commonly used in the practice of usability practitioners. Usability results that were collected and applied included findings regarding specific research questions in user studies together with corresponding recommendations for redesign, quantitative results of questionnaires, and results of checklist-based usability reviews. Information about the work tasks of users was of particular importance.

In some organizations, usability-related information was already managed actively while in others existing documentation of results was accessed in an ad-hoc manner. Participants reported that existing results were managed using different information systems, ranging from paper-based files and network drives to dedicated document servers and databases that offered the capability of full-text searching. Lists in spreadsheet documents could give an overview of existing studies and could be used to sort by different criteria such as number of users, user group, research questions, and the version of a product that was being tested [interview participant (IP) 2]. Blogs and wikis were able to help communicate new interface specifications throughout an organization, for example in order to increase consistency across products and discuss higher-level standards (IP 3).

Use Case	#P	#P scenario
UC1 Direct reuse for design decisions	6	4
UC2 Creation of internal standards	6	4
UC3 Overview of relevant usability results within a domain	7	2
UC4 Benchmarking and comparing usability	5	2
UC5 Tracking implementation of usability results	6	1
UC6 Controlling and reporting of usability activities	5	0

Table 1: Use cases for usability information: number of participants mentioning the use case in the interviews (#P) and in the scenarios chosen by the participants (#P scenario)

Participants also gave examples of integrating usability results into tools for project management and issue-tracking.

The use cases (UCs) recounted by the participants in the open question, that is, before questions about particular use cases could have introduced biases, most often involved direct reuse of information for design decisions and the creation of internal standards (Table 1).

UC1 – Direct reuse for design decisions

Participants frequently reported cases where they had to make a specific decision between several design options for an interface and searched for information on user behavior such as task descriptions and use cases (IPs 1, 6, 7): “*The type of work, of different work roles, does not really change over the years*” (IP 7). Results from user studies concerning particular interface elements could also be consulted and could be transferred from a product in a different domain (IPs 5, 8): “*When designing forms, we go back to old tests all of the time*” (IP 8). However, when reusing empirical results in this way, the specific context of use and expectations of users at that point in time had to be taken into consideration.

UC2 – Creation of internal standards

Internal usability guidelines, interface patterns, and styleguides for design, as well as heuristics and checklists for evaluation, could be derived from existing product specifications (IPs 3, 4, 6, 8) and by adapting and evaluating external guidelines based on internal empirical results (IPs 2, 5, 6, 7, 8). In the case of IP 8, internal evaluation of new guidelines with dedicated user studies was mandatory. Links could be defined to connect patterns or styleguides to the empirical results on which they were based to make them more transparent.

UC3 – Overview of relevant usability results within a domain

Most participants reported routinely collecting all available results at the beginning of a project, especially if they were new to a domain. One of the motivations for this was to

avoid conducting repeated studies on questions that had already been solved (IP 2). This search process could take up to two working days (in the case of IP 7) and involved consulting colleagues on how to obtain more information. Participants defined a domain in terms of the relevant work domain (e.g. procurement in enterprise software, IP 7) or the product type (e.g. mobile apps, IP 5).

UC4 – Benchmarking and comparing usability

In addition, collected usability results could be analyzed to perform product comparisons and to create benchmark values on the basis of quantitative results (IP 8), on the basis of user tests repeated at regular intervals (IPs 2, 6), or on the basis of internal usability checklists and guidelines (IPs 6, 7; cf. UC2). In order to contextualize results of evaluations, participants stated that the specific interfaces that had been evaluated should be documented or made available for reference (IPs 1, 8).

UC5 – Tracking implementation of usability results

Other examples involved the tracking of the implementation of usability findings using checklists (IPs 3, 4, 7) and in dedicated studies (IP 6). In addition, the rationales for design decisions could be collected (IP 4), and results from several studies and from customer support could be accumulated to reinforce the priority of user requirements (IPs 6, 7).

UC6 – Controlling and reporting of usability activities

Participants did not report examples in which the results of different usability methods were compared in order to evaluate the methods themselves. When asked directly about this use case, one participant (IP 5) remembered a minor project that had tried to accomplish this. The explanations of the participants related to the considerable resources necessary for this task. However, a related use case was addressed in the interviews: five participants (IPs 3, 4, 5, 6, 7) frequently reported the advantages and successes of usability-related activities to other stakeholders within the organization, especially to decision makers. This could involve the publication of all results within the organization (IPs 3, 4, 6) or the targeted presentation of only selected results (IPs 5, 7).

Perceived Barriers

While usability information was already being collected and applied to different use cases, participants also complained that they were often unaware of existing results obtained by other teams and departments, especially in the case of large, decentralized organizations, for example in the case of IP 2. In addition, it was possible that designers and developers who were responsible for product development initially perceived the publication of results as public criticism (IP 8). Some participants expressed their concern that published usability results could be inadvertently misunderstood or could be misused in an argumentation out of context by people outside the usability team who did not have the expertise to judge the applicability to a current situation. An important obstacle that was expressed by

Figure 1: Prototype of a usability information system with filter options in the left column (content-type, product, module, problem cause, severity rating, usage scenario) and a search box at the top of the page. Result-List (center) displays findings (*Erkenntnis*) representing positive results and potential usability problems (with the number of users, user group, and severity rating). Design solutions (*Entwurf*) are marked either with their priority or as implemented. The right column shows statistics from user experience questionnaires for the user tests currently selected, and recommends relevant user-interface guidelines (*Richtlinien*).

participants was the additional effort needed to collect and systemize results, which should therefore be motivated by the intended use cases of an organization, and, if possible, should offer advantages for the ongoing project at the same time.

FOCUS GROUPS

As the range of use cases that were identified indicates, the possibilities for using collected usability results are rich and diverse. Focus groups were conducted to gather and prioritize requirements regarding the organization of usability information to make them searchable and applicable in the use cases that had been described. One focus group was conducted with the usability team of a large software company developing enterprise software (nine participants), the other with the usability team responsible for a major online shop (six participants). Contact with both companies was established through two participants of the interview study. While providing more contextualized results, representatives of both companies also took part in the evaluation of the interface, which may have biased the results.

During the focus group, participants first shared examples of use cases of usability information in their work. They then discussed possible criteria for searching (e.g. by date, project, colleagues responsible for a study) in the context of these cases. Participants were provided with a list of possible criteria that had been compiled from the analysis of existing approaches and frameworks as well as the results of the interviews. After the discussion of the criteria, the participants were asked to prioritize them individually. A group discussion took place in between. The results demonstrated differences between the two companies regarding the approaches to the reuse of usability information.

The company developing enterprise software used a component-based approach to interface development. Therefore, the concept of an *interface element* was well defined and was used to aggregate and search for usability information. In this company, usability information was used to inform design decisions (UC1) and to establish new standards (UC2). Search criteria mentioned included the *context of use*, especially *tasks from a domain perspective*, as well as *user groups* and the *usage situation* (mobile/stationary).

In the context of the e-commerce company, aggregations of results from existing studies served as a basis for design decisions (UC1) when there was not enough time to conduct a user study. In addition, certain aspects of the software were evaluated with user studies at regular intervals, focusing on recurring research questions (UC4). Synopses of several studies were produced and combined with external results to be used to create reports for other stakeholders within the company, especially at the management level (UC6). Participants judged aspects of the *context of use*, such as the *user task* and the *user group*, as the most important criteria for accessing usability information. *Research questions of studies*, such as scrolling behavior on the homepage of the shop, were also regarded as important for accessing relevant results independently of specific products. Participants in this group also mentioned *product-related criteria*, but to a lesser degree. Prevention of innovation was mentioned as a potential negative consequence of usability information management if outdated results were used as arguments for design decisions without reflection of the current context.

DESIGN OF A USABILITY INFORMATION SYSTEM

The results of the interviews and the focus groups informed a scenario-based design process [23] of a prototypical usability information system (UIS, Figure 1). The prototype focused on the perspective of information use, that is, on the support for search, retrieval, and analysis. Personas and scenarios described current usage practices of in-house usability practitioners. These scenarios were then reformulated as activity scenarios to discuss alternative means of supporting search and analysis. This procedure helped to incorporate design as a transparent step and made it possible to define requirements regarding the organization of usability information and system support needed during the search process.

These requirements were implemented as an interactive prototype. Realistic contents are a prerequisite for an evaluation of the information structure. The prototype therefore offers access to a corpus of usability information created by students in usability classes during evaluation and redesign of different real-estate websites as test objects. The prototype presents descriptions of usability issues as the main content type. These are supplemented by observations and user comments from the studies. Design recommendations suggest solutions to many of the issues. All findings are presented as distinct information objects. However, users have access to the corresponding source documents (i.e. usability reports) if the metadata provided did not offer sufficient contextual information. Answers to a standardized user questionnaire (User Experience Questionnaire [16]) provides an example of quantitative usability results. The results of the questionnaire are aggregated dynamically to match the current search results. Internal guidelines are presented as recommended contents together with the search results.

Contents can be accessed directly via full-text search or in an exploratory manner through browsing and faceted navigation in results. Faceted navigation builds on a comprehensive sample of search criteria, which cover different dimensions of usability results related to the product, the context of use, the parameters of the study, and the classification of problems (Table 2). Typed links between objects define relations of derivation, for example, from user comments to usability issues and related design recommendations. Search results can be viewed as a result list with metadata, titles, and expandable descriptions, and could also be visualized on a timeline. The prototype was implemented using *Exhibit*, a framework well suited for prototyping faceted search applications.¹ An auto-completion feature that leverages the faceted categorization provided was added.

EVALUATION OF PROTOTYPE

The prototype was evaluated to generate insights regarding the preferences of usability practitioners in terms of the content types, metadata, and means of access that are necessary for search and analysis (RQ2) as well as to validate the completeness of the functionality (RQ3) and the acceptance of a usability information system by the potential users (RQ4).

Methods and Participants

A participative cognitive walkthrough [7] was used to evaluate the prototype and to judge the principles guiding its design. The scenario used in the walkthrough was based on the examples from the preceding studies and focused on two use cases: *direct reuse for design decisions* (UC1) and *overview of relevant usability results within a domain* (UC3). During the presentation of the scenario, participants were allowed to interact with the prototype. They commented on the use of the UIS in the context of the scenario as well as its application in their own context of work. This transfer helped to focus on conceptual aspects of the prototype. In addition, all participants answered a survey after the walkthrough: acceptance of the approach was measured using scales (*Perceived Usefulness* and *Ease of Use*) from the Technology Acceptance Model (TAM) [26]. The survey included additional items formulated on the basis of the previous studies regarding risks and expected efforts, using five-point Likert-type scales. Participants also gave their personal rankings of the five most important categories for searching and the five most important features for information interaction. All sessions were recorded (audio, video of the participant and the screen). Analysis of qualitative results primarily focused on conceptual structures to identify missing and unnecessary concepts and to judge whether concepts were represented in the interface [2].

Eleven usability practitioners from ten software development companies were recruited as participants in the context of a conference for usability professionals. Three of the

¹ <http://simile.mit.edu/>

Content Types	Relations	Facets	
Usability finding	<i>Finding</i> is based on <i>Event</i>	Product	Usage scenario
Design recommendation	<i>Finding</i> is similar to <i>Finding</i> *	Severity	Interface element
Usability guideline	<i>Recommendation</i> is based on <i>Finding</i>	Product area	Problem cause
Quantitative result	<i>Guideline</i> is based on <i>Recommendation</i>	User group	Author
Recorded event	<i>Recommendation</i> / <i>Guideline</i> is validated by <i>Finding</i> *	Date	Priority*
External resources*	<i>Finding</i> is based on <i>Event</i>	Number of users*	Implemented*
		Validated*	Device*

Table 2: Conceptual structure of a usability information system. Facets ranked by importance to participants (top to bottom) – marked concepts (*) have been added as a result of the evaluation study

sessions were conducted after the conference, either onsite or remotely using screen-sharing software. The participants represented diverse professional backgrounds, work experiences (from half a year to 15 years, with a mean of 5.6 years), and companies (ranging from 60 to 230,000 employees). Five of the eleven participants had management responsibilities. Participants were required to work in a team of usability practitioners rather than as the only one responsible for usability because in this situation collecting usability results is more beneficial to an organization.

Results

In the post-session survey, Participants rated the scenario used for the walkthrough as highly or very realistic (10 of 11 participants), which underlines the external validity of this study. Regarding use cases for usability information, controlling usability activities was added to the use case of reporting (UC6) because participants, in particular those who are managing a usability team, reported that existing results are analyzed in order to optimize allocation of resources. In addition, results helped to revise the requirements regarding information organization and means of interaction for UIM and gave insights regarding the acceptance of usability information systems.

Information Organization

Participants commented on and rated the different representations of the results and the corresponding metadata that were available in the interface. The metadata provided were considered to give sufficient contextual information to apply the results (five of eleven participants agreed; five agreed strongly). Because of this, it appears to be permissible to represent results as discrete information objects instead of only in the context of a complete report.

Examples of the dimensions of product, context of use, and interface element were all rated highly in the post-session survey. While most participants ranked the facets *product* (eight participants) and *scenario* (six participants) as most important criteria to be used for searching (Table 2), the ratings of other facets, such as *interaction element*, *severity*, *problem cause*, *user group*, and *author*, differed between participants. The *count of users* who are affected by a problem can serve as a more objective criterion of importance than a *severity rating* and in addition helps to estimate the reliability of a finding. Metadata related to the development

process, such as *priority* and *state of implementation*, can support tracking of implementation of usability findings (UC5), as well as controlling and reporting (UC6).

In addition, typed relations between the objects were favored by some of the participants over a classification. Additional relation types were proposed which represent the similarity of usability issues or denote design recommendations that have been validated empirically to solve a usability issue (Table 2). Some participants, namely evaluation participants (EPs) 1, 7, 10, 11 also stated that there was a need to derive generalized problem descriptions, which would constitute a content type of an intermediary level of granularity between the findings of a single user test and internal guidelines.

Transferability of usability results between projects was discussed in the walkthrough at a point in the scenario when a transfer of a finding between modules with similar functionality for two different products was described. Five of the participants commented on this situation. Most comments demonstrated that the similarity of both the products and the respective user scenarios need to be considered (EPs 7, 8, 10, 11). However, if findings can be traced back to general characteristics of human cognition, they are perceived as less dependent on the product and the context of use. A common cause of usability problems was another indicator that was used to judge the transferability of existing solutions (EPs 2, 9).

Information Interaction

Preferences regarding search appeared to differ on an individual level: Five of the participants rated full-text search as important while another five preferred faceted search and the exploration of results. Participants also considered browsing using links between objects to be an important feature. While participants found additional quantitative, aggregated results from questionnaires to be less relevant (EPs 1, 2, 3, 11), comments showed a positive attitude towards the application of result counts in the faceted filters, i.e. counts of qualitative results, as a means of quantitative analysis, especially those participants with management responsibilities. Most participants rated recommendations of relevant guidelines as a form of higher-level generalized results as important (eight mentions in the survey). Two participants (EPs 2, 9) commented that guide-

lines could also serve as an additional entry point to access the more detailed results, especially for those stakeholders without a background in usability. Some of the participants asked for the integration of external resources, e.g. from published usability studies (EPs 4, 7, 10, 11). A time-based view of results triggered the idea of representing relations between results in a thread-based view (EPs 8, 11).

Acceptance of a Usability Information System

The acceptance – and consequently the use – of an information system is considered to be dependent mainly on usefulness and ease of use as perceived by its users [26]. In the study, participants rated usefulness relatively highly (giving it a mean of 4.1 over four items of perceived usefulness of the TAM on a scale from 1 to 5 with a standard deviation of 0.6). There were no significant differences between the items emphasizing effectiveness (quality of deliverables) and efficiency (productivity). This outcome was confirmed by the results of additional survey items on the potential influence of a UIS on the avoidance of errors and on the realization of recommendations (quality) compared to the avoidance of unnecessary activities (efficiency). The usefulness of a UIS was also rated highly for other stakeholders in the company in a single-item question. However, only four participants stated that the results of a UIS should be accessible to everyone in an organization. Three participants apparently expected misuse of this information and three expected resistance to the publication of results from other stakeholders in an organization.

Ease of use received significantly lower ratings (mean = 3.2; standard deviation = 0.66) than usefulness, which could be expected of an experimental prototype. However, even though it was not addressed in the scenario in this study, the effort needed for data acquisition would probably have a considerable influence on the adoption of a UIS in an organization, as has been stated in previous interviews and focus groups. In the survey, participants mostly expected a high degree of effort, but at the same time many of them stated that an additional effort of 2–3% (five EPs) or more than 3% (four EPs) would be acceptable in a project.

DISCUSSION OF FINDINGS

The empirical studies presented demonstrate that usability information is already used for a variety of use cases and on a regular basis. The main contribution of this work is in the empirical collection of use cases for usability information, their validation, and the discussion of requirements of a usability information system based on a realistic prototype. The results help to advance the research area of information management for usability results. Although the methodical approach provides indications of the benefits of a structured and integrated approach to the use of existing results, the perspective is primarily that of usability practitioners as potential users. Because of this, no direct conclusions about the effectiveness and efficiency of the practices described can be drawn. Evidence for this will have to be collected in case studies of real projects.

RQ1: Use Cases

The use cases for a collection of internal usability results encompass the need to search directly for specific findings (UC1), the need for exploration and learning (UC2, UC3), and the possibilities for comparing, analyzing, and presenting results (UC4, UC5, UC6). This taxonomy of use cases was validated in the focus groups and in the evaluation study. Examples of the use cases have shown up in previous studies, but have never been described comprehensively. The use case of reporting collected results externally appeared as a new motivation to classify usability results, which is congruent with the communication needs of in-house usability practitioners [6]. For usability practitioners who are working in specialized domains, it may be of particular importance to exchange information about the domain [4].

RQ 2: Criteria for Searching

The results demonstrate that attributes describing the context of use and aspects of the product both constitute important dimensions that may be used to search for, explore, and select relevant usability results for transfer to current questions. This supports the general assumptions of the domain theory [25] and scenario-based knowledge management for usability information [9]. In addition, indicators of reliability are required to support judgments of the reliability of results. Considering the differences in the ratings of search criteria between the two organizations in the focus groups and between participants in the evaluation study, it appears to be necessary to adapt the information structure to the needs of each organization, confirming the assumptions of [27] and [8]. Differences may arise as a result of the characteristics of the products developed because of different development processes and because of preferences for different use cases of usability information. Therefore, explicit relations between similar usability issues, their design solutions, and derived generalizations may facilitate the emergence of a classification system in an organization. In combination with a higher-level faceted classification that provides metadata on the context of use and the product structure, this may avoid the need for a predefined, more detailed classification system and corresponding additional efforts necessary to classify results.

The question of how to search for usability results is closely related to the decision about which result should and can be applied in which context. When asked to judge the applicability of existing information in the context of a new project, participants appeared to assign different scopes of representativeness to different types of usability results: results that were based on basic mechanisms of human perception; for example, findings about the size of buttons were considered to have comparatively broad scope which increased their confidence to transfer them to other products. Insights about specific interaction elements, such as the design of elements on online forms, could often be transferred directly to similar features in other products. For the transfer of findings on users' behavior and their work

tasks, the similarity of the respective contexts of use had to be evaluated. While the results about users' behavior were specific to certain contexts, they were also perceived as comparatively stable over time. In addition, participants addressed several general aspects of the reliability of those results that were to be transferred: the frequency of occurrences, repeated insights from several user studies, and subjective trust in the author. Results corresponding to predefined research questions were also considered to be more reliable. This points towards the potential for hypothesis based usability testing, which can considerably improve reliability of reusing results of previous studies.

RQ 3: System Support

The results of the evaluation study demonstrate that the possibilities for exploring and analyzing results should complement direct searching in order to support awareness and interpretation of available information. Based on the characteristics of the use cases, exploration is probably useful to acquire an overview of available results (UC3) and to create new guidelines (UC2), while analysis of result sets supports comparing (UC4), tracking (UC5), and especially controlling and reporting of results (UC6). In addition to classification, the definition of relations between similar findings was described as the most important means of structuring usability information, especially for the construction and management of internal guidelines (UC2). This process should be supported based on automatic suggestions of similar results. This corresponds to the requirement to browse usability results in a threaded form in the context of other results that confirm or contradict them, which points to the need to document the iterative validation of results, a simplified form of representation of a design rationale [18].

Results may be collected and managed at different levels of aggregation and comprehensiveness. When organizations collect reports from usability studies, an overview of studies can help to identify which existing reports need to be consulted. This may be useful for learning from existing research (UC3). In addition, the most important research findings may be extracted and organized using relevant metadata. This will help to quickly find relevant results for design decisions (UC1) and to derive new usability guidelines and patterns (UC2). The comprehensive collection of all results, together with quantitative research data as demonstrated in the prototype in this study, can additionally be used to support comparative analysis and benchmarking (UC4), tracking of the implementation of usability results (if integrated with general issue management), and presentation of accomplishments of the usability department (UC6), for example, to support cost-justification. However, to support stakeholders other than usability engineers, optimized means of access should be offered.

RQ 4 – Acceptance

Judged by overall ratings of usefulness, acceptance of support for UIM can be expected to be high if requirements are

prioritized according to the importance of the relevant use cases of usability information in an organization. Participants in the evaluation study stated that they would accept the additional efforts necessary for data acquisition. In our studies, however, the benefits of applying structured results were not judged in relation to the additional efforts needed to structure these results. Participants mentioned additional factors that could decrease the motivation to share usability information. These corresponded to barriers described in the literature on knowledge sharing: the possible perception as a public form of critique may present a problem in companies that have a low tolerance for mistakes [20], while concern about opportunistic misuse of published usability information may result from a lack of trust in colleagues in large companies [20].

CONCLUSIONS AND OUTLOOK

Usability practitioners expect to see improvements in both the effectiveness and the efficiency of their work from actively managing usability information as a resource. When use cases of usability information management from empirical studies in organizations are compared to the goals of existing tools and frameworks, the gap between research and practice becomes apparent: while frameworks from research mainly provide support for a reliable diagnosis of results [1] and for communication within projects [27], there is a need to communicate and analyze results outside the scope of a single project. We expect the results presented here to provide a starting point as well as a blueprint for a process for organizations to create their own internal systems for UIM.

However, optimizing access to usability information is only one aspect of an efficient and effective use of these information resources. Additional research will be necessary to improve support for data acquisition, leveraging tools that are already used to record and communicate usability results. In addition, the results indicate that support for continuous, user-driven knowledge construction, for example, linking and grouping results, is a promising area for further research. In addition to an information systems perspective on UIM, the optimization of processes and communication structures in an organization will therefore have to be considered more thoroughly in future research.

ACKNOWLEDGMENTS

This study was made possible by the enthusiasm of all participants and by the support of the German UPA. The feedback of the reviewers considerably helped to improve this publication.

REFERENCES

1. Andre, T.S., Hartson, R. and Williges, R.C. Determining the effectiveness of the usability problem inspector: a theory-based model and tool for finding usability problems. *Human Factors* 45, 3 (2003), 455–482.
2. Blandford, A., Attfield, S. *Interacting with Information*. San Rafael, Morgan & Claypool, 2010.

3. Carroll, J.M., Singley, M.K., Rosson, M.B. Integrating theory development with design evaluation. *Behaviour & Information Technology* 11, 5 (1992), 247–255.
4. Chilana, P.K., Wobbrock, J.O., Ko, A.J. Understanding usability practices in complex domains. *SIGCHI 2010*, ACM (2010), 2337–2346. <http://doi.org/10.1145/1753326.1753678>
5. Følstad, A., Law, E.L.-C., Hornbæk, K. Analysis in practical usability evaluation: a survey study. *SIGCHI 2012*, ACM (2012), 2127–2136. <http://doi.org/10.1145/2208276.2208365>
6. Furniss, F., Blandford, A., Curzon, P. Usability work in professional website design: insights from practitioners' perspectives. In *Maturing Usability*, Law, E.L.-C., Hvannberg, E.T., Cockton, G. (eds.), Springer, London, 2008, pp. 144–167.
7. Grigoreanu, V., Mohanna, M. Informal cognitive walkthroughs (ICW): paring down and pairing up for an agile world. *SIGCHI 2013*, ACM (2013), 3093–3096. <http://doi.org/10.1145/2470654.2466421>
8. Ham, D.H. A model-based framework for classifying and diagnosing usability problems. *Cognition, Technology & Work* (2013), 1–16.
9. Haynes, S.R., Carroll, J.M., Rosson, M.B. Integrating user-centered design knowledge with scenarios. In *Human-Centered Software Engineering – Integrating Usability in the Software Development Lifecycle*, Gulliksen, J., Seffah, A. (eds.). Springer, Dordrecht, 2005, pp. 269–286.
10. Heo, J., Ham, D.H., Park, S., Song, C., Yoon, W.C. A framework for evaluating the usability of mobile phones based on multi-level, hierarchical model of usability factors. *Interacting with Computers* 21, 4 (2009), 263–275.
11. Hornbæk, K., Frøkjær, E. Comparison of techniques for matching of usability problem descriptions. *Interacting with Computers* 20, 6 (2008), 505–514.
12. Howarth, J.R., Smith-Jackson, T., Hartson, R. Supporting novice usability practitioners with usability engineering tools. *International Journal of Human and Computer Studies* 67, 6 (2009), 533–549.
13. Hughes, M. A pattern language approach to usability knowledge management. *Journal of Usability Studies* 1, 2 (2006), 76–90.
14. Hughes, M., Reeves, T. Team learning in usability testing. In *Qualitative research in technical communication*, Conklin, J. Hayhoe, G.F. (eds.). Routledge, New York, NY [u.a.], 2011, pp. 331–345.
15. International Organization for Standardization. ISO/IEC TR 25060 – Common Industry Format (CIF) for usability – General framework for usability-related information, 2010.
16. Laugwitz, B., Held, T., Schrepp, M. Construction and Evaluation of a User Experience Questionnaire. *HCI and Usability for Education and Work*, Springer, 2008, pp. 63–76.
17. Pyla, P.S., Howarth, J.R., Catanzaro, C., North, C. Vizability: a tool for usability engineering process improvement through the visualization of usability problem data. *Proc. 44th Annu. ACM Southeast Regional Conf.*, ACM (2006), 620–625. <http://doi.org/10.1145/1185448.1185584>
18. Regli, W.C., Hu, X., Atwood, M.E., Sun, W. A survey of design rationale systems: Approaches, representation, capture and retrieval. *Engineering with Computers* 16, 3 (2000), 209–235.
19. Rice, S., Thaker, J., and Wichansky, A.M. ISO 25062 usability test planning for a large enterprise applications suite. *DUXU 2011*, Springer (2011), 185–192. http://doi.org/10.1007/978-3-642-21675-6_22
20. Riege, A. Three-dozen knowledge-sharing barriers managers must consider. *Journal of Knowledge Management* 9, 3 (2005), 18–35. <http://doi.org/http://dx.doi.org/10.1108/13673270510602746>
21. Roschuni, C., Goodman, E., Agogino, A.M. Communicating actionable user research for human-centered design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 27, Special Issue 02 (2013), 143–154. <http://doi.org/10.1017/S0890060413000048>
22. Rosenbaum, R. The future of usability evaluation: increasing impact on value. In *Maturing Usability*, Law, E.L.-C., Hvannberg, E.T., Cockton, G. (eds.). Springer, London, 2008, pp. 344–378.
23. Rosson, M.B., Carroll, J.M. *Usability Engineering: Scenario-Based Development of Human–Computer Interaction*. San Francisco, Morgan Kaufmann, 2002.
24. Rosson, M.B., Maass, S., Kellogg, W.A. Designing for designers: an analysis of design practice in the real world. *SIGCHI 1987*, ACM (1987), 137–142. <http://doi.org/10.1145/29933.30873>
25. Sutcliffe, A. *The domain theory: patterns for knowledge and software reuse*. L. Erlbaum Associates, Mahwah, NJ, 2002.
26. Venkatesh, V., Bala, H. Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences* 39, 2 (2008), 273–315.
27. Vilbergsdottir, S.G., Hvannberg, E.T., Law, E.L.-C. Assessing the reliability, validity and acceptance of a classification scheme of usability problems (CUP). *Journal of Systems and Software* 87 (2014), 18–37.
28. Woolrych, A., Hornbæk, K., Frøkjær, E., Cockton, G. Ingredients and meals rather than recipes: a proposal for research that does not treat usability evaluation methods as indivisible wholes. *International Journal of Human–Computer Interaction* 27, 10 (2011), 940–970. <http://doi.org/10.1080/10447318.2011.555314>