An Implementation of User-Experience-Based Evaluation to Achieve Transparency in the Usage and Design of Information Artifacts

Margit Scholl
UAS Wildaup, D-15745 Wildau, Germany
margit.scholl@th-wildau.de

Abstract
User experience (UX) is a concept that covers the total effect of all the elements of an IT system on the user. An information artifact (IA) can contain any kind of information. In order to increase transparency in the usage and design of online and offline IA, the TEDS framework was implemented in Moodle as an electronic evaluation activity (TEDS*MOODLE) to evaluate the Moodle course rooms. Moreover, for the sake of enhancing UX transparency, it was extensively tested with students from the UAS Wildau’s administration and law department from September 2013 to January 2014 and used to evaluate two online training courses and other kinds of IA. The results demonstrate the functionality and limitations of the application for different kinds of IA. Individual UX can be used to improve the quality and design of IA. The application itself is based on DIN EN ISO 9241-110 and confirmed the recommendations of this standard.

1 Introduction and Background

The term “artifact” comes from the Latin: arte stands for “by art or craft” and factum for “something made.” An artifact is therefore an entity that is generated by human beings, and an “information artifact” (IA) can contain any kind of information. It may be a book, newspaper, TV ad, website, document, a special service, or an information system in its entirety. Online information systems (oIS), e-Government services in general, and learning management systems (LMS), in particular their content, can be seen as IA. As Taylor showed in 1982 [1], context- and environment-neutral IA are barely capable of supporting the needs of human actors. IA need to be designed with a clear understanding of their relative value so that they can deliver this value to human actors (users) in a particular context. The specific context could be geographical, organizational, social, intellectual, or cultural.

Within a wide range of disciplines worldwide, learning processes are nowadays increasingly supported by the use of oIS and LMS, and these systems are being developed to facilitate the action-oriented and self-directed learning of users [2]. However, use intensity driven by technology alone does not necessarily lead to user acceptance and the expected knowledge outcomes. The design of these systems and the didactic scenarios they contain are becoming increasingly important, for example, LMS designed with e-portfolios and taxonomies to address the complex challenges of modern teaching and learning processes [3] [4] [5] [6]. This is closely connected with the learning system’s quality assurance, which should focus on the concrete needs, the individual competences, and the specific usage contexts of the human actors involved. From such a perspective, technical systems and learning platforms as instruments of human–computer interaction (HCI) are becoming instruments of intentional social interventions (cf. [7]). As Jeff Johnson (2014) has pointed out, HCI, seen as a three-level interaction of engineering, design, and science, is not easy, because the phenomena to be exploited in HCI range from the abstractions of computer science to psychological theories of human cognition, perception, and movement [8], which is why user-interface design guidelines are developed in practice. However, to properly understand them, a theory of design that analyzes and describes them in more detail is necessary. Gregor & Jones (2007), for example, focus on the structural components of design theories to establish design knowledge as theory [9].

Moreover, the development of oIS and LMS should be done “with” the users. In order to gain insight into the quality of a platform, it should be analyzed, together with the human actors, using a clearly structured approach that gives insights into the need for different usage scenarios and the variety of human actors involved. Such an approach is adopted by Taylor’s Value-Added Model [1] and design ideas that have been further developed and transformed to create the so-called TEDS framework [10].
view to achieving greater insight through systematic evaluation and increasing acceptance of oIS/LMS in our education system, the TEDS framework was electronically integrated into our Moodle learning environment at the University of Applied Sciences (UAS) Wildau as part of the TEDS@wildau project, which was funded by the European Regional Development Fund (ERDF) and the university [11]. The evaluation application activity, now fully developed and integrated, is called TEDS*MOODLE. It is a purpose-customized electronic system solution based on the TEDS framework and should, on the one hand (aspect #1), be used electronically to evaluate other IA like websites, TV advertising, or books. On the other hand (aspect #2), this evaluation application is itself an IA and should be developed, designed, and discussed as a design science research methodology (DSRM), as shown in [12]. This paper tries to clarify both aspects.

TEDS*MOODLE was fully operational in the winter semester 2013–14, complete with didactical and technical support. It was in place in various courses, with forty students from the administration and law department participating in each case: it was tested for the Moodle course rooms in parallel with lectures and tutorials (see [13]), and for two online (expert) training courses, the one covering intercultural competences and the other, the possibilities of electronic collaboration, as shown here. In addition, some student groups with their own projects were to select other information artifacts and apply the TEDS framework and the application TEDS*MOODLE themselves. Other students were asked to think about the individual criteria, their meaningfulness, and their limitations. However, the student evaluations, both of the Moodle course rooms and the online learning courses with the TEDS*MOODLE activity, were voluntary and, as a result, had varied participation rates. The methodology was also used in other student projects to test TEDS*MOODLE for different kinds of IA and to promote the sustainability of this application. This was supplemented by student papers designed to suggest new ideas for ways to improve this application activity.

Based on the TEDS framework policy of discussing the assessment results with the users as evaluators [10], we were able to make improvements not only to the IA we examined (aspect #1) but also to TEDS*MOODLE itself (aspect #2). The empirical results show how individual informational offerings can be evaluated in a user-oriented fashion and also give concrete suggestions on ways to improve the IA in question. The specific research questions (RQ) in this practice part of the project are as follows:

RQ#1 What general lessons can be drawn from the evaluation results for the tool? (aspect #1)
RQ#2 What is the overall impact of applying the TEDS framework in an electronic environment, in particular as the purpose-customized electronic system solution TEDS*MOODLE? (aspect #1)
RQ#3 How can the TEDS*MOODLE activity be made self-descriptive and easy to use for different IA evaluations? (aspect #2)
RQ#4 How can we make it possible for inexperienced users (evaluators) to adequately understand and differentiate the categories and criteria of the evaluation without further didactic and technical support? (aspect #2)
RQ#5 What standard can be used to assess the practical performance of the tools, and what conclusions can be drawn for their further development? (aspect #2)

Section 2 provides literature reviews for the TEDS framework, describes the implementation of the TEDS*MOODLE application activity, and gives earlier results and further background information. Section 3 deals with the case studies involving the evaluation of two online training courses as well as the results of the student projects based on different kinds of IA evaluations. The individual results of the case studies for the TEDS framework and the methodological approach are summarized in the first part of section 4. In the second part, the current state of the TEDS*MOODLE evaluation application is benchmarked against the requirements of DIN EN ISO 9241-110:2008-09 to estimate its practical utility. Section 4 also provides answers to the research questions. In section 5, the conclusions and an outlook for the continued use of this sophisticated evaluation tool is presented.

2 Literature Review and Earlier Results

As Hassenzahl & Tractinsky (2006) have pointed out, UX has become a buzzword in the field of HCI and interaction design, and interactive products are now not only more useful and usable but also fashionable and fascinating objects of desire [14]. The use of technology for learning or working tasks facilitates and heavily influences new forms of activities, social rules, and interactions (see [15]). The range of interactive possibilities can be extremely diverse [16]. However, machines, systems, technologies, software, and services do “act” [see 7]—just how they act is an important question in the process of developing services for human target groups [17].

---

1 The staff on the TEDS@wildau project, which finished on 6/30/2014, were Peter Ehrlich and Denis Edich, who shared a job.
Moreover, elements of UX have become a vital reference for the World Wide Web (WWW) and international interaction designers and have come to define the core principles of practice-oriented developments [18].

The use of learning platforms in Blended Learning scenarios and the extension of traditional learning approaches to include constructivist methods have also raised other concerns for the learning process, the user experience, and the learning outcomes (cf. [19] [20] [21]). Meanwhile, learners are typed, using various categorizations for different types of e-learning users: discovery-oriented, research-oriented, network-oriented, and education-oriented users (cf. [22]). Based on such findings, attempts have been made to diversify learning platforms and their contents with more precision and adapt them to the different individual needs of teachers and learners. The technique itself—seen as a key component of electronically controlled teaching and learning—is of such a view to a “didactic actor,” which changes the behavior of both teachers and learners in principle, influenced and channeled (see [15]). Therefore, evaluations of such IA from a user perspective are particularly important if we are to make the actual use and effect of the design more transparent.

This is where the TEDS*MOODLE application can be used electronically in a similar experimental way, as an implementation of UX to improve transparency and gather evaluation data on the usage of IA. It is based on the TEDS framework [10], which itself has taken Taylor’s categories [1] [23] a great deal further, presenting a finely structured, analytical instrument for evaluating IA with a clear methodology and specific emphasis on actors and usage scenarios. The TEDS-framework approach is amenable to any kind of IA, comprises a series of thirteen-step procedures [10], includes forty evaluation criteria, and can be summed up as follows (see [11]):

- Identification of *personae* and their concrete wants, needs, values, and beliefs.
- Identification of specific *scenarios* of utilization as hypothesized archetypes of human action in context.
- Evaluation of an information artifact in a clearly defined thirteen-step procedure.
- Discussion and production of detailed recommendations for designing and redesigning the information artifact in question.

The TEDS-framework methodology has been used exclusively in the last two years to evaluate the websites of professional sports teams [24] [25] and in the analysis of the teams’ social media and mobile IA [26]. These results demonstrated the general functionality of the TEDS-framework methodology and generated recommendations for sports managers for fan acquisition from an information management perspective. In addition, the TEDS-framework methodology also inspired research investigation, as shown in [27]. However, the TEDS*MOODLE evaluation application is the first electronically based implementation of the TEDS-framework methodology. The TEDS*MOODLE integration concept and interfaces are shown in more detail in [11].

The original idea for the development of TEDS*MOODLE was to create a tool based on the scientifically accepted TEDS-framework method and perform continuous evaluations of the Moodle learning platform worldwide, with particular emphasis on our own Moodle classrooms and learning content. The evaluation has been implemented with multilingual capabilities for future target groups in English, Spanish, and German in order to make the process more comprehensible.

The development process and the integration concept of TEDS*MOODLE were shaped by the original research questions: 1) How can we implement the analytical power of the TEDS framework into an effective application on different devices for the user-specific quality assurance of learning scenarios? and 2) What didactic and technical support is needed to establish TEDS*MOODLE as a continuously used evaluation method to enhance electronic environments? (see [11]). In order to answer these questions, TEDS*MOODLE was built as an activity in Moodle, while also running as an independent, modular tool. Another goal was to offer users of the evaluation method a simple computer-aided guided tour of the evaluation criteria.

![Figure 1. When using TEDS*MOODLE the evaluators can see the individual criteria as follows: the top line indicates the main category and sub-criterion, followed by an explanatory question and the Likert rating scale; on the right a so-called key frame is shown.](image)
However, the forty criteria are not self-explanatory for laypeople; users have had to be trained at some length in the methodology in order to be able to evaluate an IA using the Likert scale without misunderstandings occurring.

To obviate a long process of instruction, the individual criteria were originally assigned two questions each as a way to clarify their meaning. Based on the students’ user experience, the questions were further clarified and limited to one each [13]. Because the individual evaluation criteria with their “subtle differences” are complex and open to interpretation, both the reorganized definitions and the questions are easily available in the system. To give future users (evaluators) simple and self-explanatory access to the methodology and the categories used (see fig. 1), as well as an easy understanding of the process, these specific user questions have been developed for each criterion (for the list of criteria and questions (see [13] in English, [28] in German).

A further idea was to illustrate the individual criteria with a short video or a so-called key frame (see fig. 1). The videos were discarded to cut down on the evaluation time. The pictures were evaluated by staff from different projects and are currently assigned uniformly in all three language versions, although the linguistic–cultural component of the images needs to be considered. The TEDS*MOODLE tool, however, has been designed so that both the questions and the picture (key frame) can be changed in a simple way.

At the beginning of the TEDS*MOODLE development process, a further research question came up: Should all existing categories and sub-categories of the original TEDS framework be used unchanged and, if not, which should be modified or removed for our specific electronic learning environment? [11] Although the original forty categories of the TEDS framework are not meant to be used arbitrarily, it is still possible that the procedural steps, in particular, may need to be adjusted to the categories, because not all evaluation criteria can be used for the specific Moodle evaluation. For example, because of the specific corporate design applied by the UAS Wildau and the general functionalities of Moodle, it might be useful to adapt the evaluation criteria for practical efficiency. As part of this specific implementation, some of the categories were given particular weighting and considered in more detail, while other evaluation criteria were seen as less relevant in the context of Moodle. For this specific integration we use thirty-three from the original forty evaluation criteria in a flexible implementation that can be changed if necessary. Following the original TEDS framework, the TEDS*MOODLE application activity distinguishes the following six main evaluation categories [13]:

- Ease of Use (criteria 1–5),
- Noise Reduction (criteria 6–12),
- Quality (criteria 13–17),
- Adaptability (criteria 18–26),
- Additional Performance Features (criteria 27–28), and
- Affection (criteria 29–33).

All the TEDS*MOODLE criteria are set up according to a German-style Likert scale: “Strongly agree” (1), “Agree” (2), “Neither agree nor disagree” (3), “Disagree” (4), and “Strongly disagree” (5). The difference between the American and German Likert scales is simply the assignment of the points awarded in the evaluation, which now run automatically in TEDS*MOODLE. In the American (and Spanish) system the best score is 5 and the worst 1. In the German system, it is exactly the opposite: the best score is a 1. If the evaluators want to interpret their results themselves and process them, then this difference should be taken into account. Going through all thirty-three criteria, one arrives at a summary of all the evaluations (see fig. 2). Here, as long as the rating is left open by the instructor and remains visible as a Moodle activity, the evaluator can still make and save changes at any time.

**Figure 2. Excerpt (the first six sub-criteria) from the full list of thirty-three criteria at the end of an IA evaluation using the TEDS*MOODLE application.**

Put another way, TEDS*MOODLE is an analytical tool for formulating design specifications that has the ability to ensure acceptance on a person-to-person...
basis and makes provision for user-oriented quality management. The very first results of the implemented evaluation of our Moodle homepage are shown in [11]. The results of three Moodle course room evaluations are summarized and discussed in [13]. In this paper, the focus is on the benefits of applying the evaluation tool for any form of IA.

Taking an IA development in the context of design science (DS) theory [9] and research methodology, we should ensure that a DS research process complies with various guidelines [29] and includes several steps, such as problem identification and motivation, definition of the objectives of a particular solution, design and development, demonstration, evaluation, and communication [12]. Moreover, in the TEDS*MOODLE case studies shown here, there are two aspects to consider: first, the usage of the TEDS framework methodology for an evaluation of other IA and, second, the improvement and design of the tool itself. So, on the one hand, we need to interpret the results of the following sections relative to the TEDS-framework methodology (aspect #1) and, on the other, we should derive recommendations for designing the tool using DSRM (aspect #2). In the DSRM context, all the student projects described in the following sections can also be seen as the artifact evaluation of TEDS*MOODLE itself.

3. Case Studies

The adapted TEDS framework contains the following steps (see [2] [3] [17]): the first steps are to determine what should be evaluated as an IA and to find a reference IA, an “anchor,” which is identified either as outstanding (a positive anchor) or poor (a negative anchor); a further key step is to identify the specific scenario and persona, i.e. the target group in contextualized human activity. Those things are defined when the TEDS-MOODLE application activity is launched. Each survey is performed separately with this tool. First, the evaluation criteria are applied for the anchor and the results are discussed by the evaluators/voters (target group), and then the interviewing process is carried out for the IA. In order to identify variances and assessment discrepancies, all the voter evaluations are statistically analyzed. The current tool version creates several evaluation-data files in table form (raw data, Open Calculator, and Microsoft Excel) with statistical data for each criterion and an initial graphical analysis based on the anonymous statistical data of the evaluators. After each methodological step, there were extensive conversations between students and lecturer, to clarify misunderstandings and introduce quality control (aspect # 1), and to generate ideas for developing the design of the evaluation tools (aspect # 2). However, the lecturer did not actively intervene in the results so as to keep the process objective and to see how well the students understand the TEDS methodology and are able to apply the TEDS-MOODLE tool. Moreover, it was of interest to see if the decision to have the application implemented in other projects—for other IA and by other actors—would work. The assessment of student project results could also be carried out using anonymous raw data, as was used in the evaluation of this report.

3.1. Evaluation of Two Online Training Courses

Independently of the TEDS@wildau project, two online training courses were developed for staff in SMEs as part of the InterKomp KMU 2.0 project. The purpose of these courses is to give SME staff, as well as trainers, a basis for developing sensitivity to intercultural problems in a modern work environment [30]. Both online courses are available in German and offered for free to guests and Wildau students in the Moodle learning platform.

As part of the TEDS@wildau project, evaluations using the TEDS-framework methodology were then submitted by students in the administration and law program. The intercultural course was used as (positive) anchor and the IA was the online course e-Collaboration. After each evaluation, lecturers and students discussed any variances, thus consolidating the basis for comparison. An iterative research process was applied, whereby the questions developed within the TEDS-MOODLE application were used to clarify the purpose of the individual assessment categories. Each evaluation was carried out using a German-style Likert scale (see above). These evaluations were also statistically analyzed and any variances were pointed out. Due to the voluntary nature of the process, there was a difference in the number of students evaluating the two courses: the anchor evaluation of the online training course Skills for Intercultural Work had thirty-one voters/students (see fig. 3), compared with only nine for the IA, the online training course e-Collaboration (not shown).

Despite the different levels of participation, the structuring of the TEDS-MOODLE assessment was very similar: the negative “rogue results” (which scored 4 and 5) are for the most part more or less identical. This is surprising, because three years of experience went into constructing the anchor course,

2 http://kmu-interkomp20.th-wildau.de/?page_id=356
http://kmu-interkomp20.th-wildau.de/?page_id=313
which is very different from the IA, e-Collaboration, which was the first course we established. All in all, with an average of “Agree” (2) and a variance close to 1, the criterion Satisfaction (#33) performs rather worse for the anchor than expected—even though the value for the IA is worse still, with an average of “Neither agree nor disagree” (3) and a variance of 0.44. However, this might change if the target group really needed such information in daily life. For both courses, many criteria achieved on average a good ranking with “Strongly agree” (1) and “Agree” (2). Nevertheless, even a seemingly straightforward category like Ease of Use shows unexpected potential for improvement in criteria #1 and #2: it is evident from this that Browsability (#1) and Presentation (#2) need to be enhanced. Moreover, even the anchor has some less than positive assessments in Aesthetics (#29) and Entertainment (#30), whereas Engagement (#31) gets a better ranking. Neither course managed the good scores that were hoped for in Stimulation (#32).

Contrary to the intentions of the course developers, there are more significant differences in the evaluation of the criteria #10, 19, 20, 21, 22, 24, and 26: Precision of Search (#10) and the main category Adaptability (#18–26) need to be improved. However, as discussed in [13], #24 (Individualization/Adaptability)—which is characterized by the question “Can you tailor the way information is presented to meet your individual requirements?”—has among the worst ratings. This is not surprising since, in essence, students were not provided with any self-directed activities relating directly to Adaptability as part of rights management in the Moodle course room—the administrators of the Moodle platform do not allow this. The worst performers are Transaction (#21) and Privacy Policy (#26), because there is no direct transition function available for purchasing goods and products, and little information about how to secure and protect user data. For many criteria, a large variance or standard deviation from the mean exists, reflecting the individuality of the UX.

3.2. Tool Application in Student Projects
3.2.1. Motivation of the Student Projects in the Administration and Law Program. Students of administration and law are required to complete a project assignment on e-Government in their fifth semester. At the beginning of the semester, they are given an introduction to the methodology of the TEDS framework and the TEDS*MOODLE application activity; they are required to use the university’s Moodle learning platform, and the language of communication is German. Their task involves conducting a new TEDS evaluation of their own with TEDS*MOODLE. Altogether seven project teams were formed with between four and seven participants in each team; they were able to choose the new application scenario, the anchor, the IA, and the target group, and were required to document the process.

Team 1 asked whether wissen.de is, in fact, a better online knowledge portal than Wikipedia (see table 1). The aim in comparing the two encyclopedias was to determine which of the two web pages best complies with the desires and expectations of the student team. All the project teams took part in identifying the strengths and weaknesses of the two sites and were asked to submit suggestions for improvements where necessary. This involved the TEDS*MOODLE application activity, its methodology, and the catalogue of questions developed by the real-life project TEDS@wildau [11], which was here used in its German version [28].

The field of tourism played a key role in team 2’s choice of a local authority website, as this represents...
an important source of income for the two administrative districts (Dahme-Spreewald and Oder-Spree), which are local recreation areas for the Berlin metropolitan region.

Table 1. Evaluations performed by student project teams (#1–7) using the TEDS-MOODLE application, and their choices of a specific anchor and IA. Students of administration and law, 5th semester.

<table>
<thead>
<tr>
<th>PT #</th>
<th>Anchor</th>
<th>Information Artefact</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Internet presence for administrative district Dahme-Spreewald, Tourism</td>
<td>Internet presence for administrative district Oder-Spree, Tourism</td>
</tr>
<tr>
<td>3</td>
<td>News website ARD Tagesschau, Homepage</td>
<td>News website RTL Now, Homepage</td>
</tr>
<tr>
<td>4</td>
<td>Menu &lt;Papa Nò&gt;</td>
<td>Menu &lt;Vapiano&gt;</td>
</tr>
<tr>
<td>5</td>
<td>Homepage &lt;City of Magdeburg&gt;</td>
<td>Homepage &lt;City of Potsdam&gt;</td>
</tr>
<tr>
<td>6</td>
<td>Online complaints/ Online suggestions City of Potsdam</td>
<td>Online complaints/ Online suggestions City of Düsseldorf</td>
</tr>
<tr>
<td>7</td>
<td>Main “offers” pages Deichmann</td>
<td>Main “offers” pages Tamaris</td>
</tr>
</tbody>
</table>

Team 3’s documentation showed the results of a comparison of the web pages of two German online news services—the public television station ARD and the private broadcaster RTL. ARD’s daily news program, Tagesschau, was defined as the anchor, because at first sight this website was deemed particularly good. Besides the seriousness of the information presented, the decisive factor was the station’s treatment of the news and its level of trustworthiness. RTL Now was used as the comparison website—its design, where news is accompanied by advertising, was seen as questionable.

Team 4 looked at a completely different IA: menus as information carriers. This led to discussions about the adaptation and interpretation of the categories and explanatory questions for the two cases under examination, as there were a number of “differences of interpretation” in the evaluation of the anchor’s individual categories. Since the project was conducted “in situ,” there was subsequent discussion about the need for correction measures to separate the conditions in the restaurants where the assessments were carried out from the evaluation of the menus. In this kind of situation, the students could have used their smart phones, as the tool has been programmed for different mobile devices [11].

Team 5 took the TEDS-MOODLE survey system as a tool to examine the extent to which the guidelines for Germany’s e-Government strategy have been implemented in selected authorities. As the student projects are required to restrict their focus to two public administrative bodies, they are not qualified to assess the quality of e-Government in other German authorities—the results should be focused on the scenario and the persona that have been defined. The personal experiences of the team members showed that the City of Potsdam administration’s main strength lies in the direct contact it has with members of the public in its offices. But they also note that, by contrast, the City of Potsdam’s Internet presence lacks clarity and falls short in its attempt to appeal to users. The anchor role is played by the homepage of the City of Magdeburg, because its Internet presence is seen as particularly “good” and the city has certain characteristics that it shares with Potsdam.

In contrast, team 6 chose the City of Potsdam’s Internet presence as a negative anchor, in particular in terms of the online complaints and suggestions coming from citizens. The team took the City of Düsseldorf as the IA. The key question was “Which authority offers the best online complaint service based on the TEDS-MOODLE evaluation procedure?”

Team 7 focused on the area of e-commerce and chose Deichmann’s online shoe shop as the anchor and Tamaris’s as the IA.

3.2.2. Results of Student Projects. A vigorous discussion of all the categories took place within and between the project teams. The aim was to consolidate the basis for comparison so as to establish the best benchmarks for evaluating the individual IA. As part of these discussions, subjective perceptions and interpretations were exchanged to facilitate better understanding of the categories (aspect #1). For the real-life project TEDS@wildau [11], this provided the basis for reworking the German questions, which were reduced to one concise question per criterion (aspect #2). One of the main discussion points that came up with all the teams was the neutral (“3”) evaluation, which was frequently used when the evaluators were uncertain whether the particular functionality they were being asked about even existed. Another outcome of the discussions was the overall consolidation of the anchor and IA evaluations.

Team 1’s evaluation showed that the greatest potential variances are in the main categories of Adaptability, Additional Performance Features, and Affection. One reason for this was thought to be the subjective perception of the individual categories. The results of their evaluation using the TEDS-MOODLE methodology showed that in five of the
six main categories the *Wikipedia* anchor fared better than wissen.de, which only achieved a more positive evaluation in the area of Quality. However, this is a key factor in research and knowledge, and the result is influenced by the fact that wissenmedia’s publishing program, which operates the Internet platform wissen.de, incorporates series of encyclopedias, like Bertelsmann or Brockhaus, as well as Faksimile. But while, from an academic point of view, *Wikipedia* is not considered “quotable,” the students of this group focused more on simplicity and the speed with which information is displayed and accessed, and definitely see *Wikipedia* as the more practical digital reference tool, although not for research purposes.

In its discussion of the individual evaluation categories, team 2 exposed the problems inherent in the main category Affection and, more specifically, the sub-category Aesthetics. Since evaluations here are primarily governed by emotional factors, discussion will not necessarily lead to conformity in the individual assessments. This being the case, the results will remain extremely varied and will not be characterized by “compromise.” For one of the administrative service providers, this can only mean offering different websites for different target groups, corresponding to a finely judged definition of *persona* and *scenarios* and the sophisticated results of the evaluations.

Team 3’s comparisons produced suggestions for improving both news sites (ARD and RTL) in terms of individual customization and the introduction of different language preferences. In both cases, it was thought that the privacy notices needed revising with respect to encryption and legitimate expectation. The numerous advertisements on *RTL Now*, which are positioned amongst the news content, were considered a nuisance that drew attention away from the news. Here it should be stressed that an evaluation of this kind also has a cultural component. Performing this evaluation in different countries might well lead to very different results and varying appraisals, which means that websites should be tailored to the specific country they are intended for.

Team 4, which by and large gave a more positive rating to the anchor’s menu than to the IA, concluded that the evaluations were also influenced by individual taste preferences: while one of the evaluators liked sushi, the other preferred Italian cuisine—this also had an impact on their assessments of the menus. By extension, this discussion relates to the need to be able to customize and individualize IA.

Team 5 proceeded from the assumption that citizens are mainly concerned with issues surrounding public services, which are part of every local authority’s self-managed administrative tasks. Every citizen wants to know which official to take their enquiries to, or at least which department they need to go to. The upshot of this is that a government Internet site is expected to present the structure of the administration so that it is clear and can be quickly called up; it should also appeal to the user and have navigable topic areas. It was not possible to resolve all the disagreements in the evaluations in the individual criteria. But for the *City of Potsdam* there were a number of suggestions for improvement: while an *average user* of the homepage is probably familiar with it as an informational source and is satisfied with the information provided and the data protection functions, the team found a weakness in the web presence in the following categories: Mediation, Precision, Community, Aesthetics, and Stimulation. All in all, it is evident that, despite the similarities between the two cities selected, their home pages differ in terms of Quality. To return to the team’s original question, it was ascertained that the administrations of the two cities have implemented the guidelines of Germany’s e-Government strategy, although the standard of implementation clearly differs. In this way the TEDS*MOODLE* evaluation can be used to establish best practices.

The *City of Potsdam* uses the application *Maerker* to manage complaints and suggestions—this is a successful e-Government project put in place by the State of Brandenburg. However, in particular for the final category and the question “Do you have a sense of satisfaction after retrieving information?” team 6 gave the city a uniformly poor score. Their dissatisfaction stems from the fact that with the “traffic light” display you can only see whether or not a complaint is being processed, while the result is not shared with citizens online. This fuels our motivation to develop TEDS*MOODLE*. It also reinforces the idea that, with regard to the details of how communication with administrative services is handled, the expectations of citizens (particularly young people) differ from those of the administration concerned.

Team 7’s results indicate that they had some difficulty evaluating a few of the TEDS categories. The category Novelty caused particular problems since the companies’ Internet sites gave nothing to go on here. In addition, it was only after placing an order that you found out whether the particular product was available. The team thus recommended incorporating the availability (of products) in the TEDS*MOODLE* evaluation scheme. In Germany, privacy is another important factor to consider. Some team members were unable to find the privacy policy on the website, even though it was posted. It was located right at the bottom of the page in a much smaller font size and was thus easy to miss. The team suggested using larger lettering or making it more prominent, for ex-
ample in a pop-up window. Feedback and Community were also evaluated as being subpar. The option to use the “contact” feature to send an email or obtain the address and telephone number was not viewed as sufficient feedback. This must be given a much more active role for customers/consumers. Community was nonexistent both for the anchor and the IA, and site operators should remedy this.

4. Findings for the TEDS*MOODLE evaluation
4.1. Summary of the case studies concerning RQ#1 and RQ#2 (aspect #1)

It should be noted that the TEDS-framework methodology [10] always refers to a very specific scenario and is defined with a very specific, uniform target group that will be included as evaluators in the assessment of IA. Therefore, the results can only be generalized up to a point, because they always relate to the defined scenario and target group, or, to use Taylor’s model [1, 23], UX is of relative value in personalizing and diversifying information systems and IA. Of course, the bigger the uniform target group in the defined scenario is, the more the results can be generalized. For small groups of up to ten evaluators, different results from a similar group cannot be excluded, and the statistical evidence seems to be weak in correlating data and conclusions. However, UX is always about the specific needs of this target group, so that even with smaller audiences first indications of possible changes in design and usability of the investigated IA are quite remarkable.

The weighting of the evaluation categories is another aspect that should be considered. In the discussions the student teams highlighted a number of criteria and questions that they see as most important for their anchor/IA. Actually it is not always clear how to rate all the criteria for all kinds of IA. More importantly, the evaluation categories and can be interpreted according to the selected anchor and IA. This interpretation requires a broad understanding of the categories and can be interpreted by the evaluators as a limitation of the method or the tool.

These findings should be reflected in further scientific studies on the practical design guidelines of Jeff Johnson, in particular with respect to his remarks on “learning from experience” and the “factors affecting learning” [8].

We come now to the overall impact of applying the TEDS framework in an electronic environment, particularly as the purpose-customized electronic application TEDS*MOODLE (RQ#2). One aspect is its simple accessibility for various disciplines and in a protected environment. But if the application is to be used in other scientific investigations extending to users via the Internet, it should be implemented independently of Moodle. The other aspect is the learning outcome involving information systems using the TEDS framework: after working with the TEDS*MOODLE application and the TEDS methodology, the majority of administration and law students discovered for themselves that in future they would take a much more sophisticated view of unfamiliar media.

4.2. Findings for the TEDS*MOODLE tool: RQ#3, RQ#4, and RQ#5 (aspect #2)

To summarize the findings for aspect #2 of the investigation on TEDS*MOODLE as they relate to RQ#3, the implementation is to be self-descriptive and easy to use for different anchors and IA evaluations—at the beginning of the TEDS survey, the initiator (teacher, lecturer, scholar, project leader) specifically names the evaluation (anchor or IA) and provides publicly visible details of the scenario and the target group. In addition, the programming of the tool is also aimed at mobile devices. Nevertheless, the tests showed that for non-web-based IA the tool
should be used independently of Moodle. This change from integration to separation relates to the need to clarify the new security requirements.

The changes to the (German) evaluation questions prompted by the discussions with the students that were made by the TEDS@wildau team [11] brought greater clarity for all actors. Although we have adjusted the English and Spanish questions accordingly, the author still thinks that the real-world use of the TEDS*MOODLE application activity in these languages will lead to necessary cultural improvements being made to the evaluation tool. This has to be done in further research projects. However, the key questions are one answer to RQ#4—how can we make it possible for inexperienced users (evaluators) to adequately understand and differentiate the categories and criteria of the evaluation without further didactic and technical support? Moreover the TEDS@wildau project team expected ideas to emerge from this to help them visually supplement the refined (German) key questions for each criterion with an image or video, a so-called key frame, for future evaluators. This means including key frames alongside the key questions, although this is a difficult process and there is a cultural component that requires further research, too.

To answer RQ#5, one must first benchmark the TEDS*MOODLE evaluation application against the requirements of DIN EN ISO 9241-110:2008-09 to estimate its utility based on a regulatory DIN framework. The characteristics of developed technologies, such as interactive oIS or LMS for societal processes, should be based on the general objectives of dialogue design outlined in DIN EN ISO 9241-110:2008-09 [11]:

- Appropriateness of tasks
- Self-descriptiveness
- Expectation conformity
- Learning encouragement
- Controllability
- Margin of error
- Individualization.

These objectives [31], analogous to the criteria of the TEDS framework [10], cannot be simply applied but must be specified and defined. What does each objective of dialogue design mean in the concrete context, for the specific user scenario and for the target group under consideration? Overall, from these results and for general usage of the TEDS*MOODLE tool, the recommendations of DIN EN ISO 9241-110:2008-09 [11] require the following:

- An interactive system is task-appropriate when it helps the user to do his or her job duties and when functionality and dialogue are based on the characteristics of the task rather than on the technology used for task completion [11:8]. The TEDS*MOODLE tool is made task-appropriate for the criteria with key questions.
- A dialogue between human and system is capable of self-description to the extent to which it is obvious to the user at any time which dialogue they are in, where in the dialogue they are, what actions can be taken, and how they can be executed [11:10]. TEDS*MOODLE is self-descriptive via the user interface.
- A human–system dialogue is expected to be compliant if it addresses the predictable issues and complies with generally accepted conventions. The principle of consistency increases the predictability of a dialogue [11:11]. TEDS*MOODLE ensures consistency for the anchor and the IA both in processing the criteria and in common usage.
- The dialogue is conducive to learning when it supports the user in learning how to use the interactive system and guides [11:12]. The interactive system should allow the user to perform the work task with minimal learning effort by facilitating the dialogue with minimal informational input, but with additional information available upon request [11:13]. With the general user interface, TEDS*MOODLE supports the Likert selection with minimal learning effort in the analysis options.
- A dialogue is controllable when the user is able to start the dialogue flow and influence its direction and speed until the goal is reached. The users should have control over how the dialogue will continue. He or she should be able to determine the resumption point for the continuation of the dialogue if the work item allows it [11:13]. TEDS*MOODLE ensures controllability by making it possible for users to save their evaluations and change their result.
- A dialogue is error-tolerant if the intended results can be achieved despite evident errors in input with either no or minimal correction effort on the part of the user. Error tolerance is here addressed with error detection and error prevention (damage limitation), or with error correction or error management, ensuring that any errors that may occur are properly dealt with [11:14]. TEDS*MOODLE assures this by allowing for variability in the evaluation result and incorporating fault tolerance. The graphical analysis and evaluation files are accessible to every user and can be used to promote discussion and error management.
- A dialogue can be individualized if users can change the human–system interaction and the
presentation of information to adapt to their individual abilities and needs [11:15]. TEDS*MOODLE guarantees individuality by establishing the personae, scenarios, reference context, and language of evaluation. TEDS*MOODLE’s finely structured categories and criteria provide concrete evidence that performs a normative function. As described in DIN EN ISO 9241-110:2008-09 [11:8], the dialogue principles are not strictly independent of each other and may overlap in content. It may be necessary to balance between principles in order to optimize usability. The applicability and relevance of each principle depend on the particular context of use (scenarios), on the user groups (personae), and the selected dialogue technology (IA). Using TEDS*MOODLE, the evaluation results are UX-based, achieve transparency for the usage and design of the IA under scrutiny, and suggest ways to improve the IA. However, in response to RQ#5, compromises will be made in practice in the way an IA, interactive system, or online service is designed. Nevertheless, the future development of TEDS*MOODLE in accordance with DIN EN ISO 9241-110:2008-09 is a practical step.

5. Conclusions and Outlook

The original idea [11] was to develop and implement an electronic tool specifically for LMS, but also more generally for any oIS, the aim being to enable users of this evaluation method to run a simple but scientific based computer-supported process of human-centered evaluation. Because, as a first step, we wanted to improve our Moodle platform and the various course rooms, we have implemented the analytical instrument and procedure of the TEDS framework in an application that can be called up as an “activity” in the Moodle learning platform. However, it works independently from Moodle and may be used in any oIS. In future TEDS*MOODLE development, we will need to progress from integration to separation.

The TEDS*MOODLE integration concept and interfaces are shown in [11]. For this particular LMS integration we use only thirty-three of the original forty evaluation criteria [see 13]; however the implementation is flexible and can be changed if necessary. Moreover, key questions and key frames as well as multilingual capabilities are implemented for further cultural research and can be easily adapted. This also applies to the use of the TEDS*MOODLE application activity itself: it should be tested by both purely English-speaking and purely Spanish-speaking groups to generate ideas for further cultural adaptation. Such adaptation is a necessary part of making the TEDS*MOODLE application activity available for free usage without further support.

TEDS*MOODLE thus enables the design of any IA to be assessed and improved on an ongoing basis. The underlying assumption is that the increasingly important evaluative challenges can be met and that key recommendations for IA interaction, software development, technical didactics, online moderation, and user support can be derived from UX evaluation results. From the many results and discussions, it is clear that the information provided in IA should be simpler and much better organized and capable of being easily found by the specific user groups, while feedback possibilities should be much more active. Moreover, with TEDS*MOODLE the long-standing demand for participative software design methods can be specifically addressed.

TEDS*MOODLE seems to have a fairly broad application, even beyond the educational sector. Like the TEDS framework, the TEDS*MOODLE integration that is derived from it is not limited to the academic sector and should be viewed simply as a preliminary area of concrete application. The results demonstrate the functionality of the TEDS*MOODLE application for many kinds of IA. All in all, the results reflect the singular quality of the evaluation and the contribution of individual UX.

6. Literature


