

U-HAUL Usability Investigation and Assessment

Executive Summary

This report is based on the assessment of five video recordings of user performing typical tasks that might be carried out on the U-Haul website. Users were asked to complete seven tasks:



- Rent a truck for a cross-country move
- Check on the requirements for a driver's license
- Determine the cost of appropriate an indoor storage unit
- Locate a store phone number
- Add specific boxes and packing resources to an existing truck rental
- Determine damage liability
- Locate a U-Haul store closest to a specific location

The test activities were carried out 25 – 28 March 2011 in a small office/meeting room. A moderator provided some background information to each participant and managed the test sequence from an adjacent room. Participants were encouraged to talk-aloud about their decisions as they progressed through each task. The moderator had participants estimate task difficulty (1 – 7 scale) before each task and then assess ease of use and confidence (Assessed Success) after the completion of these tasks.

Task #	Participant Task success	Assessed Success		Time on Task		Estimated Difficulty		Assessed Difficulty		Difficulty Divergence	
		Min	Max	Min	Max	Min	Max	Min	Max	Better	Worse
1 - Truck Rental	5	4	7	77	686	4	6	2	7	3	2
2 - License	5	6	7	25	64	2	6	5	7	5	0
3 - Storage Unit	4	1	7	138	363	5	6	1	7	1	3
4 - Phone #	4	6	7	63	116	5	7	6	7	2	0
5 - Additions	5	3	6	139	542	4	7	2	6	0	4
6 - Damage	2	3	6	106	170	1	6	2	4	1	0
7 - Location	5	6	7	48	97	6	7	6	7	1	1

Sample sizes are small so only the ranges of task times for successful completion are reported. All five users found that task #2 (identification of license requirements) to be easier than expected – all located this information by going to the FAQ links. Four of the five users found the task of adding boxes and packing supplies to an existing truck rental to be more difficult than expected. Only two of the five participants were correctly able to determine

responsibilities for damage (LDW) and most indicated they would require further assistance to validate their understanding.

The five recorded sessions were reviewed in detail by the evaluator. The examination yielded a number of findings (positive and negative) and these have been rated on their impact and categorized using a set of 15 heuristics. This approach allows for summary information and recommendations to be presented. The following table indicates the number of problems or positive findings identified.

Rating	Rating Code	Description	# Identified
Critical Problem	A	Causes frequent catastrophes.	0
Serious Problem	B	Delays test participants in their use of the website for some minutes, but eventually allows them to continue.	3
Minor Problem	C	Causes test participants to hesitate for some seconds	74
Good Idea	I	A suggestion from a test participant that could lead to a significant improvement of the user experience.	21
Positive Finding	P	This approach is recommendable and should be preserved.	1
Bug	X	The website works in a way that is not in accordance with the design specification. This includes spelling errors, dead links, scripting errors, etc.	2

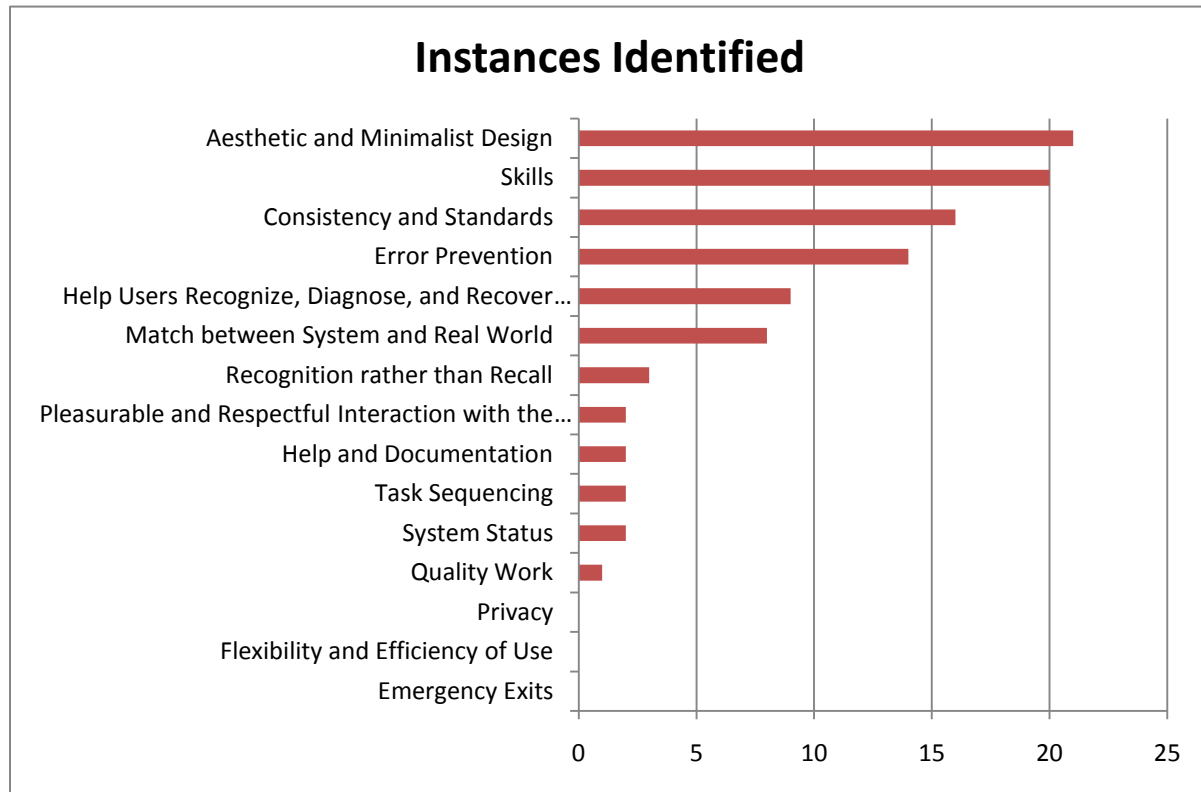
Three serious problems were identified in the evaluation. These in some way could influence users to ignore information or be confused in a significant manner. Most problems (or findings) were of a minor nature and were typically found in the heuristic categories associated with *Consistency & Relevancy*, *Task & Work Support* and *Error Recognition & Recovery*. Positive findings have also been included as participants often commented on features or support they found useful.

Two bugs were identified and these might require further investigation as they both resulted from slow presentation of images and could be specific to bandwidth, browser or system performance of the users' computers.

Recommendations

Users sometimes got confused or made incorrect assumptions about the information presented as they completed tasks. On some pages, there was a significant volume of information presented and this contributed to information being missed. Some inconsistency in the display of content also probably contributed to this confusion. A reduction in

the volume of information would likely improve the decision making of users, result in faster completion times for tasks, and less inaccuracy in order s (less in-person support would result). Further work should be done on the presentation of content on insurance coverage.



An A/B study might be an effective way to examine the impact of making aesthetic or content reduction changes to U-Haul task-related pages. Click through or drop-off rates could provide data for appropriate statistical analysis on task completion of alternative designs.

Participants were able to locate the FAQ materials and certainly scanned items at the top of the list. The FAQ might be made more visible. None of the participants located questions/answers regarding Damage Coverage. This might be because of the long list of Questions. A reorganization of this document such as with a quick access table might make content easier to locate.

Background to study

A usability assessment of U-Haul web-based tasks is to be carried out by examining five recorded user sessions. These task-based activities ensure that user-specific actions are analyzed and the decision-making strategies of users are incorporated into the assessment. Usability issues (Tullis & Albert, 2008) could be:

- Anything that prevents task completion
- Anything that takes someone 'off-course'
- Anything that creates some level of confusion
- Anything that produces an error
- Not seeing something that should be noticed
- Assuming something is correct when it is not
- Assuming a task is completed when it is not
- Performing the wrong action
- Misinterpreting some piece of content
- Not understanding the navigation

The analysis of video recordings provide a way of determining task success, task timings and assessment of user interface interactions. A moderator ensured that users were consistently prepared for the test and could step-in should some guidance or clarification be needed. The test activities were carried out 25 – 28 March 2011

User Profiles

The background information on the five test participants is provided in the following table.

Participant	Duration (min)	Sex-Age	Occupation	Web Savvy
M1	32	F-35	Quality assurance	Average
M2	47	M-25	Education	Average
M3	42	F-35	Customer Experience	High
M4	26	M-35	Librarian	High
M5	38	M-45	Project manager	High

Description of Test Protocol

This was a moderated test activity which took place in a small office/meeting room. The user environment was pre-defined with a browser appropriately set to a starting location. A PowerPoint presentation provided some context for

the test activity and sequenced the user through tasks in a specified order. The switch between the PowerPoint Presentation and Web interaction was managed by the moderator from a remote workstation.

Test Orientation

Participants were asked if they had moved within the past five years and if so how they had arranged for this. This was done to help provide some reference or context for the participants, in part to prepare them for the nature of the test activity and possibly to exclude anyone who had completed such an activity recently.

Once a participant was “ready to start” the moderator left the room and went to an adjacent area to monitor the test remotely. The moderator confirmed with each participant that they could both be heard and understood before the test began. All participants appeared very comfortable communicating with the moderator and there was no perceived delay in responses between the participants and the moderator.

Participants were provided with a paper copy of the test activities and were able to refer to these at any time. To initiate the test activities and to ensure comprehension of each task, participants were asked to read-aloud each task before starting that task. Once the task was read, the website was displayed and the participant was in control of any computer actions. Participants talked aloud as they progressed through the assigned tasks.

Moderator Behaviour

The moderator was consistent in the treatment of all participants. She was clear and precise in communicating actions. She seamlessly took control of the participant’s computer to alternate between PowerPoint and Internet browsers.

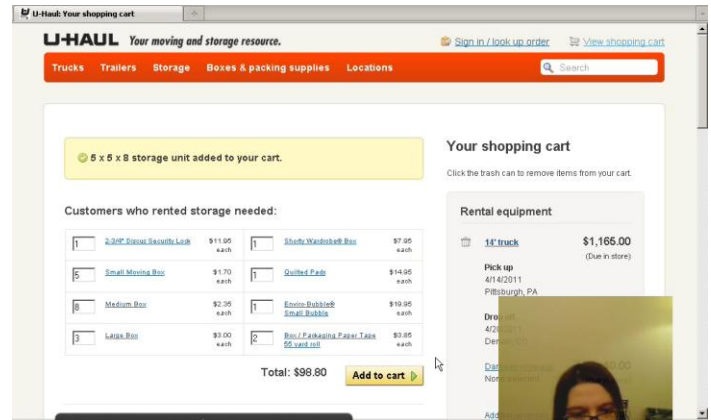
Before the start of each task, participants were asked to read-aloud the task and to indicate the expected difficulty (ease of use, 1 – 7 scale) of each task. Once a task was deemed by the moderator to be finished, the participant was presented with a PowerPoint question on the Assessed Difficulty of the task (1 – 7 scale) and a question on confidence in accuracy (1 – 7 scale). If a task was not judged to be completed successfully, the moderator skipped either or both of these post-task questions.

The moderator limited engagement with test participants once the testing commenced. The moderator, on occasion, encouraged the participants to qualify a response, reassess their position/decision/response or identify that they were satisfied with actions and were ready to proceed.

Test Activities

Users were asked to complete seven tasks:

1. Rent a truck for a cross-country move
2. Check on the requirements for a driver's license
3. Determine the cost of appropriate an indoor storage unit
4. Locate a store phone number
5. Add specific boxes and packing resources to an existing truck rental
6. Determine damage liability
7. Locate a U-Haul store closest to a specific location



Results of tests

The following table summarizes the task success, task performance and participant responses to pre-task and post-task questions on task difficulty.

Task #	Participant Task success	Assessed Success		Time on Task		Estimated Difficulty		Assessed Difficulty		Difficulty Divergence	
		Min	Max	Min	Max	Min	Max	Min	Max	Better	Worse
1 - Truck Rental	5	4	7	77	686	4	6	2	7	3	2
2 - License	5	6	7	25	64	2	6	5	7	5	0
3 - Storage Unit	4	1	7	138	363	5	6	1	7	1	3
4 - Phone #	4	6	7	63	116	5	7	6	7	2	0
5 - Additions	5	3	6	139	542	4	7	2	6	0	4
6 - Damage	2	3	6	106	170	1	6	2	4	1	0
7 - Location	5	6	7	48	97	6	7	6	7	1	1

Task Performance

Tasks #1, #2, #5 and #7 were successfully completed by all participants.

Tasks #3 and #4 were successfully completed by four of the five participants.

Task #6 was successfully completed by only two participants.

Participants confident in their activities typically scored their assessed success a 6 or 7 in the post-test question. Lower scores were in the 1 to 3 range if participants struggled with a task. In some instances, participants were encouraged by the moderator to review their response or to reexamine a task objective to complete the activity. This direction was only suggested and not directive.

Task Times

With such a small sample size, some caution must be taken if interpreting average user timings from these test activities. The range of values is probably a more valuable indicator of the effort users are prepared to put into the tasks presented. Task times were determined from the video recordings.

Tasks #2, #4 and #7 were the fastest tasks – these tasks all required participants to find a fact (locate specific information).

Activity- based tasks, task #1, #3 and #5 had the widest range of times amongst the participants. The widest range was found in task #1. Participants were persistent in their efforts to complete these tasks.

Task #6 is a judgment task (participants had to make an assessment) and had a range of times from 106 – 170 seconds. This time interval might be influenced by participant's prior knowledge or reluctance to continue searching for more information without additional support.

Task Difficulty

Participants predicted task difficulty prior to the start of each task and then assessed the actual difficulty upon successful task completion.

All five participants found that task #2 (identification of license requirements) to be easier than expected – all located this information by going to the FAQ links.

Four participants found the task of adding boxes and packing supplies to an existing truck rental to be more difficult than expected.

Only two of the five participants were correctly able to determine responsibilities for damage (LDW) and most indicated they would require further assistance to validate their understanding. Three users anticipated that this task would be particularly difficult. Only one participant scored the task easier than anticipated at the completion of the task.

Participant Motivation

All participants seemed motivated to complete the tasks presented. The scenarios were real-enough for them to want to complete. Participants were thorough in their approach and appeared to score their efforts consistently.

Heuristic Evaluation

The five recorded sessions were reviewed in detail by the evaluator. The examination yielded a number of findings (positive and negative) and these have been rated on their impact and categorized using a set of 15 heuristics. This approach allows for summary information and recommendations to be presented. The following table indicates the number of problems or positive findings identified.

Rating	Rating Code	Description	# Identified
Critical Problem	A	Causes frequent catastrophes.	0
Serious Problem	B	Delays test participants in their use of the website for some minutes, but eventually allows them to continue.	3
Minor Problem	C	Causes test participants to hesitate for some seconds	74
Good Idea	I	A suggestion from a test participant that could lead to a significant improvement of the user experience.	21
Positive Finding	P	This approach is recommendable and should be preserved.	1
Bug	X	The website works in a way that is not in accordance with the design specification. This includes spelling errors, dead links, scripting errors, etc.	2

Three serious problems were identified in the evaluation. These in some way could influence users to ignore information or be confused in a significant manner. Most problems (or findings) were of a minor nature and were typically found in the heuristic categories associated with *Consistency & Relevancy*, *Task & Work Support* and *Error Recognition & Recovery*. Positive findings have also been included as participants often commented on features or support they found useful. An alternative presentation of issues and findings by heuristic is presented at the end of this report ([Heuristic classification of issue/error/observation](#)).

Recommendations

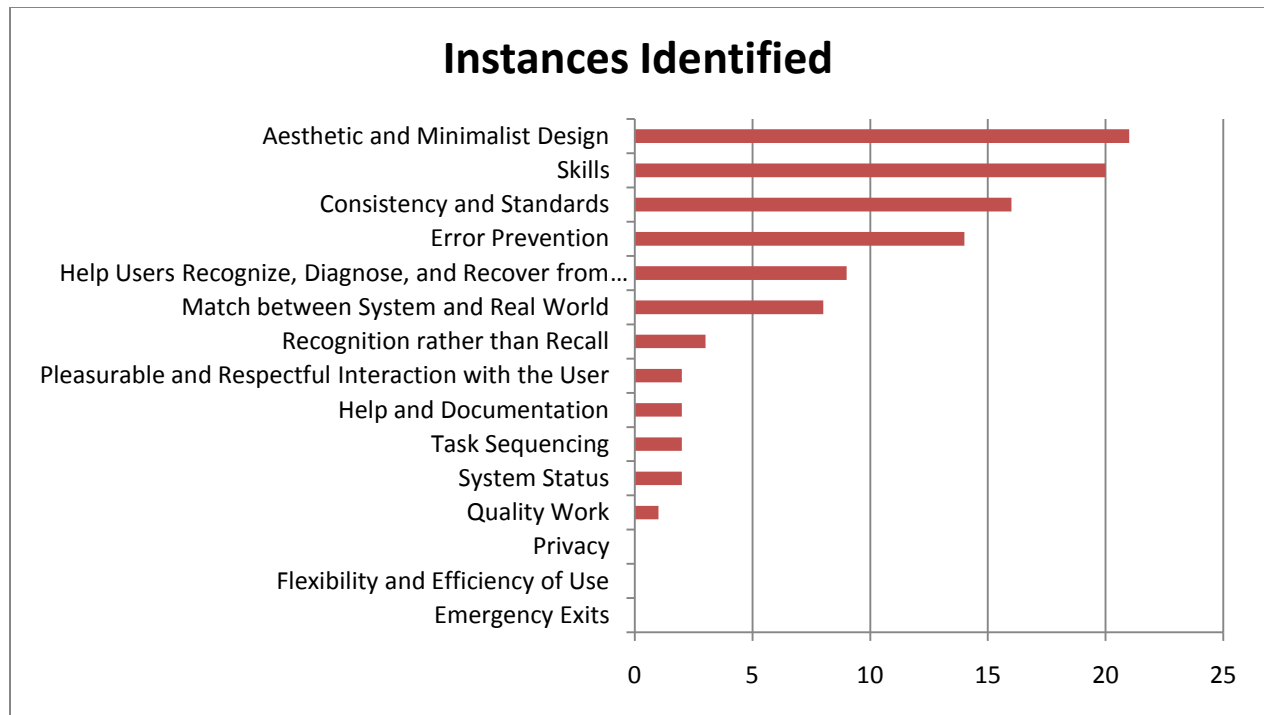
Losing the scent of information

Jared Spool of UIE talks about users losing the scent of information as they progress through a web site or application. Here, some of the participants got off-track as they moved back and forth between pages, or by not recognizing where information was presented and making quick assumptions regarding content. Where tasks are specific with a clear goal (such as renting a truck), participants would invest the time to review actions, locate mistakes and recognize and end point. This table provides summary assessment of each task and a general recommendation.

Task Description	Results	Observations	Recommendation
1 Book a truck for a long-distance move	Most successful task	Participants were familiar with the idea of choice of locations close to the home address. There was an expectation of display by price and distance. Participants decided on truck size based on image, accompanying text and prior knowledge.	Participants appear to quickly scan though the content on vehicle insurance. A comparison table might help decision making.
2 Driver's license	Most successful task. Fastest task overall.	Participants quickly looked for this information within the FAQ	FAQ could be easier to locate. Reduce the number of information links in the Quick Links at the bottom of the page.
3 Indoor storage unit	Mostly successful task	Participants associated this task with moving and were comfortable looking for the nearest locations.	Some confusion regarding the relationship of these facilities to U-Haul. Some inconsistency in rates displayed.
4 U-Haul phone number at nearest pick-up location	Mostly successful task	Perhaps the order of the tasks confused one participant as a storage location was chosen.	Clearly distinguish between U-Haul rental and storage locations.

Task Description	Results	Observations	Recommendation
5 Additional moving and packing supplies	Most successful task.	Almost all participants were confused by what was in their shopping carts. Sub-optimal performance. Participants typically found this task to be more difficult than expected.	Participants gave positive acknowledgement for suggested items, but were also frustrated by default quantities pre-assigned. There were some difficulties in determining whether items were added to the cart or not. Immediate and pickup costs should be presented side-by-side as those below the fold (below the scroll region of the screen) are easily missed. Examine the arrangement of optional materials and the expected sequence of the task. Too many options for users early in a task activity increases the likelihood of error.
6 Damage costs	Least successful task.	Participants were not confident in answer and would typically seek further support (contact). Participants were not prepared to invest a lot of time to investigate.	Continue to refine the content regarding insurance coverage. Possible have an Insurance FAQ or comparison table that supports judgment questions.
7 Nearest location at another address	Mostly successful.	Some learning/familiarity with this type of task is demonstrated in the narrower range of task times.	Possible distinguish between rental and storage locations using different icons on maps. Could allow a user to recover from a selection error quicker.

Usability issues and findings were identified and classified according to 15 heuristic categories proposed by Mueller, Matheson, Page and Gallup (1998). The figure below clearly shows that most issues fall under the categories dealing with the presentation of content (aesthetics & design, supporting user skills, consistency, and error prevention).



Heuristic classification of issue/error/observation

Category	Description	Heuristic Code	# Identified
System Status	System Status	1	2
User Control and Freedom	Task Sequencing	2	2
	Emergency Exits	3	0
	Flexibility and Efficiency of Use	4	0
Consistency and Relevancy	Match between System and Real World	5	8
	Consistency and Standards	6	16
	Recognition rather than Recall	7	3
	Aesthetic and Minimalist Design	8	21
	Help and Documentation	9	2
Error Recognition and Recovery	Help Users Recognize, Diagnose, and Recover from Errors	10	9
	Error Prevention	11	14
Task and Work Support	Skills	12	20
	Pleasurable and Respectful Interaction with the User	13	2
	Quality Work	14	1
	Privacy	15	0

Appendix: Participatory heuristic evaluation

A set of heuristics were used to categorize usability issues and findings. This set of heuristics is based on a list of 15 categories identified by Muller, Matheson, Page and Gallup (1998).

M.J. Muller, L. Matheson, C. Page and R. Gallup, Participatory heuristic evaluation, Interactions 5 (5) (1998), pp. 13–18.

System Status

1 SYSTEM STATUS. The system keeps users informed about what is going on through appropriate feedback within a reasonable time.

User Control and Freedom

2 TASK SEQUENCING. Users can select and sequence tasks (when appropriate), rather than the system taking control of the users' actions. Wizards are available but are optional and under user control.

3 EMERGENCY EXITS. Users can

Easily find "emergency exits" if they choose system functions by mistake (emergency exits allow the user to leave the unwanted state without having to go through an extended dialogue).

Make their own decisions (with clear information and feedback) regarding the costs of exiting current work.

Access undo and redo operations.

4 FLEXIBILITY AND EFFICIENCY OF USE. Accelerators are available to experts, but are unseen by the novice. Users are able to tailor frequent actions. Alternative means of access and operation are available for users who differ from the "average" user (e.g., in physical or cognitive ability, culture, language, etc.).

Consistency and Relevancy

5 MATCH BETWEEN SYSTEM AND THE REAL WORLD. The system speaks the users' language, with words, phrases, and concepts familiar to the user, rather than system-oriented terms. Messages are based on the users' real world, making information appear in a natural and logical order.

6 CONSISTENCY AND STANDARDS. Each word, phrase, or image in the design is used consistently, with a single meaning. Each interface object or computer operation is always referred to using the same consistent word, phrase, or image. Follow the conventions of the delivery system or platform.

7 RECOGNITION RATHER THAN RECALL. Objects, actions, and options are visible. The user does not have to remember information from one part of the dialogue to another. Instructions for use of the system are visible or easily retrievable whenever appropriate.

8 AESTHETIC AND MINIMALIST DESIGN. Dialogs do not contain information that is irrelevant or rarely needed (extra information in a dialog competes with the relevant units of information and diminishes their relative visibility).

9 HELP AND DOCUMENTATION. The system is intuitive and can be used for the most common tasks without documentation. Where needed, documentation is easy to search, supports a user task, lists concrete steps to be carried out, and is sized appropriately to the users' task. Large documents are supplemented with multiple means of finding their contents (tables of contents, indexes, searches, etc.).

Error Recognition and Recovery

10 HELP USERS RECOGNIZE, DIAGNOSE, AND RECOVER FROM ERRORS. Error messages precisely indicate the problem and constructively suggest a solution. They are expressed in plain (users') language (no codes). Users are not blamed for the error.

11 ERROR PREVENTION. Even better than good error messages is a careful design that prevents a problem from occurring in the first place. Users' "errors" are anticipated, and the system treats the "error" as either a valid input or an ambiguous input to be clarified.

Task and Work Support

12 SKILLS. The system supports, extends, supplements, or enhances the user's skills, background knowledge, and expertise. The system does not replace them. Wizards support, extend, or execute decisions made by users.

13 PLEASURABLE AND RESPECTFUL INTERACTION WITH THE USER. The user's interactions with the system enhance the quality of her or his experience. The user is treated with respect. The design reflects the user's professional role, personal identity, or intention. The design is aesthetically pleasing— with an appropriate balance of artistic as well as functional value.

14 QUALITY WORK. The system supports the user in delivering quality work to her or his clients (if appropriate). Attributes of quality work include timeliness, accuracy, aesthetic appeal, and appropriate levels of completeness.

15 PRIVACY. The system helps the user to protect personal or private information—belonging to the user or to his or clients.