iTree: Skill-building User-centered Clarification Consultation Interfaces

Martina Freiberg and Frank Puppe
Department of Artificial Intelligence and Applied Informatics, Institute of Computer Science
University of Würzburg, Am Hubland, D-97074 Würzburg, Germany

Keywords: Knowledge-based Systems, Clarification Consultation, UI Design, Skill-building UI, Usability Evaluation.

Abstract: Developing web-based, knowledge-based systems (wKBS) still challenges developers, mostly due to the inherent complexity of the overall task. The increased focus on knowledge-base development/evaluation and consequent neglect of UI/interaction design and usability evaluation raises the need for a tailored wKBS development tool, leveraging the overall task while specifically supporting the latter activities. As an example for such a tool, we introduce the wKBS development tool ProKEt. With the help of that tool, we developed the novel UI concept *interactive clarification tree (iTree)* with skill-building ability, that specifically is suitable for clarification consultation systems. Also, we report a recent case study, where *iTree* was implemented for knowledge-based clarification consultation in the legal domain.

1 INTRODUCTION

Despite increasing distribution in many domains, web-based knowledge-based systems (wKBS) still challenge developers: Development of appropriate knowledge bases alone is an effortful task in terms of time and money; thus, intentional UI/interaction design and usability evaluation activities remain rather neglected. Yet, wKBS are often applied in critical or specialized contexts—e.g., consultation in the medical or legal domain—where the chosen UI/interaction style can contribute strongly to either the success or the failure of the system. Thus, UI/interaction design and usability evaluation should rather be a key factor for wKBS development. This increases the need for a tailored software tool that fosters experimentation and evaluation of novel wKBS styles. We propose the tailored software tool ProKEt, that supports efficient affordable,UI-/interaction design-focussed wKBS development while at the same time seamlessly integrating usability evaluation functionality. To the best of our knowledge there does not exist previous work to date regarding similar tools.

Regarding consultation in contexts such as the legal or medical domain, it often can be valuable to not only have *general consultation systems* available that derive one or more diagnoses based on the user input, but additionally to have specialized *clarification systems*, for investigating only one distinct diagnosis—potentially pre-selected by general consultation systems or by the users themselves. In this paper, we introduce the *interactive clarification tree (iTree)* as a novel, hierarchical clarification UI/interaction style that we developed with the help of the tool *ProKEt*; iTree thereby is particularly suitable for a mixed, diverse user population and additionally provides skill-building ability. A first study in the course of a current project in the domain of legal consultation suggests general benefits of the proposed *iTree* UI style.

Related Work. With regards to general KBS/wKBS development there exist various tailored software tools—such as JavaDON (Tomic et al., 2006), or KnowWE (Baumeister et al., 2011)—and methodologies—e.g., MIKE (Angele et al., 1998), or CommonKADS (Schreiber et al., 2001). However, such approaches still mostly focus on the design and evaluation of the knowledge base; in contrast, we propose *ProKEt* as tailored wKBS development tool that seamlessly couples efficient, agile wKBS development, creative experimentation regarding KBS UI/interaction design, and semi-automated usability evaluation activities. ProKEt can be further seen as *user-centered prototyping tool* for wKBS—a concept defined by (Leichtenstern and André, 2010) as an all-in-one tool solution for enabling efficient, effective and satisfactorily design, evaluation and analysis of developed artifacts.

Probably due to the numerous benefits of web-based systems—e.g., availability, acceptance, or maintainability—to date an increasing number of knowledge-based/expert systems seems to be devel-

Freiberg M. and Puppe F..
iTree: Skill-building User-centered Clarification Consultation Interfaces.
DOI: 10.5220/0004093701370142
In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development* (KEOD-2012), pages 137-142
ISBN: 978-989-8565-30-3

oped for the web—i.e., integrated in websites or as separate, complex web applications; recent examples are (Patil et al., 2009) or (Rahimi et al., 2007). However, such wKBS apparently are being developed in a rather ad hoc manner, not following systematic methods or processes, and not (re)using (neither providing) any patterns or best practices especially regarding the UI/interaction design—probably due to a general lack of scientific research in web-based expert systems, cf. (Duan et al., 2005). Similarly, wKBS developers seem to be individuals, performing all tasks required for developing and distributing a wKBS by themselves. This further increases the need for a tailored tool that not only renders overall wKBS development an efficient, pragmatic task, but equally important specifically supports design and experimentation with web-based UI/interaction forms and their usability evaluation.

Paper Structure. The rest of the paper is organized as follows: In Section 2, we shortly introduce the tailored wKBS development tool ProKEt. Afterwards, we discuss *iTree*, a novel hierarchical UI concept for knowledge-based *clarification consultation systems* with skill-building ability in Section 3. We report on a recent case study in Section 4, where the proposed UI style was practically implemented for a wKBS in the legal domain. We conclude with a short summary of the presented research and an outlook to prospective future work in Section 5.

2 ProKEt

ProKEt is a tailored, **Pro**totyping and **K**nowledge systems **E**ngineering **t**ool for web-based, knowledge-based systems (wKBS), that additionally provides integrated support for various usability evaluation related activities. Thereby, ProKEt specifically supports web-based *consultation* and *documentation systems*, which can be developed equally well as (pure) prototypical demo systems and as fully-fledged systems for productive use. Thereby, *extensible prototyping* is put into action, facilitating a nearly effortless transition from prototype to productive system; for a more extensive introduction of the agile, extensible prototyping and engineering process with ProKEt, see (Freiberg et al., 2012). The main application logic is implemented in Java. The resulting artifacts are Servlet-based web applications, using HTML, StringTemplate, and CSS for UI creation, and JavaScript for interactivity. Regarding the knowledge representation, an XML-based specification is used for the pure prototypes, which can be directly cerated/edited with ProKEt itself. For productive systems, *d3web* (URL d3web, 2012) knowledge bases are integrated and (mostly) replace the XML specification; the latter, however, can not directly be edited with ProKEt, thus in that case an external d3web-supporting tool such as KnowWE (Baumeister et al., 2011) is required. Yet recently, we implemented a mechanism to couple KnowWE and ProKEt artifacts, thus drastically improving and easing the workflow of UI/front-end development, knowledge base development and their integration into a productive wKBS.

For supporting the straightforward evaluation of its artifacts, ProKEt further allows for seamlessly integrating both *qualitative and quantitative* data collection both for prototypes and productive wKBS; this enables developers to assess the current development state in a favorable way at any time by conducting manifold, potentially iterative, evaluations. For *qualitative data* collection, ProKEt supports both the integration of form-based questionnaires/surveys—standards such as the SUS (Brooke, 1996) and the NasaTLX (Hart, 2006) are supported out of the box, but tailored own questionnaires can be added with no effort—and of anytime feedback—mechanisms for collecting free user feedback at any time during a wKBS session. Regarding *quantitative data*, ProKEt provides a tailored, mouse click and keyboard event logging mechanism that records all relevant actions during wKBS sessions. Based on that data, ProKEt furthermore automatically can calculate a bunch of known usability metrics—such as *Success Rate*, or *Average Task Time*—proposed e.g. by (Constantine and Lockwood, 1999), but it is equally well possible to just export qualitative and quantitative data into a standard CSV format for further investigation with external tools, e.g., standard spreadsheet calculation or advanced statistical software. A more detailed introduction of that usability extension of ProKEt can be found in (Freiberg and Puppe, 2012).

3 iTree FOR CLARIFICATION

ProKEt particularly supports the development of *consultation-* and *documentation systems*. A consultation system thereby provides decision support in a particular domain based on given user input, whereas a documentation system contrastingly focusses on supporting uniform, efficient and high quality data entry. In this paper, we propose the *interactive clarification tree (iTree)* UI style specifically for *clarification systems* as a sub-class of consultation systems.

3.1 Clarification Consultation

As a subarea of classification, *clarification* relates to hypothesize-and-test as follows: Separate, general *multiplex consultation* systems can be applied first for narrowing the complete set of potential diagnoses/hypotheses down to one or several most suitable elements (hypothesize step); each of those hypotheses can then be further investigated by a corresponding *clarification module* (test step). As shortcut, users could alternatively start directly with a clarification system for a chosen hypothesis themselves.

3.2 iTree: Skill-building Clarification

We propose *iTree* as a novel UI with skill-building ability that fosters an efficient and usable user experience in the context of clarification systems.

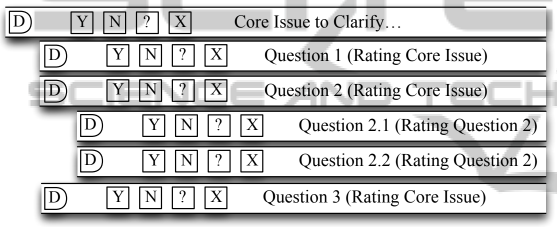

Figure 1: Schematic drawing of the *iTree* UI style.

Figure 1 presents a schematic drawing of *iTree* for clarification systems. The core issue to be rated is presented as root element of the hierarchical tree structure (Figure 1, *Core Issue*). Its rating is derived from the ratings of any desired number of top-level questions, placed directly underneath the core issue (Figure 1, *Quest.1, Quest.2, Quest.3*). Questions are a tailored form of yes/no questions with an additional value *Neutral/Uncertain*; provided answers further can be withdrawn/adapted at any time, indicated by the *X* button. The current implementation—a practical example of which is depicted in Figure 2—allows three possible abstract ratings for the core issue as well as for all questions: *Confirmed*, *uncertain/neutral* and *rejected*, which correspond to the answers *Yes*, *?*, and *No* per default. Some domains may require to swap that mapping for particular questions in favor of a more understandable question wording. Figure 2 depicts an example: The core issue is confirmed (rating: yes) if *the cancellation was NOT prohibited due to time limitations*; in that case, the swapped yes/no mapping allows for rewording the question as depicted, which is much clearer than its negated alternative. In case the user cannot answer a question directly, more detailed refinement questions—if available—can be retrieved for the current element, represented by the *D* button in the scheme and by the arrow in Figure 2; as example, the second top-level question in Figure 2 contains two refining questions, which list in more detail the conditions which confirm/reject its parent question. Question ratings are always propagated from inner levels of the hierarchy up to the topmost question(s) by either *AND* or *OR* connections. Let *pn* be a parent node and cn_p a child node of *pn*; for calculating the rating of *pn*, the following rules apply:

AND nodes:

$IF\ \exists cn_p = no\ THEN pn = no$

$IF\ \forall cn_p = yes\ THEN pn = yes$

$IF\ \forall cn_p = rated\ AND\ \exists\ cn_p = neutral\ AND\ \nexists cn_p = no$
$THEN\ pn = neutral$

$IF\ \exists cn_p = unknown\ AND\ \exists\ cn_p = neutral\ AND \nexists cn_p = no$
$THEN\ pn = unknown$

OR nodes:

$IF\ \exists cn_p = yes\ THEN pn = yes$

$IF\ \forall cn_p = no\ THEN pn = no$

$IF\ \forall cn_p = rated\ AND\ \exists\ cn_p = neutral\ AND\ \nexists cn_p = yes$
$THEN\ pn = neutral$

$IF\ \exists cn_p = unknown\ AND\ \exists cn_p = neutral\ AND\ \nexists cn_p = yes$
$THEN\ pn = unknown$

This means, e.g., that the core issue in Figure 2 is rated *yes* only if all of its children are rated *yes*, as those are connected to the core issue by AND (second rule above); likewise, *cancellation prohibited due to time limitation* is rated *yes* as soon as one of its children is rated *yes* (fifth rule above) due to the OR connection. One advantage of *iTree* is the suitability for

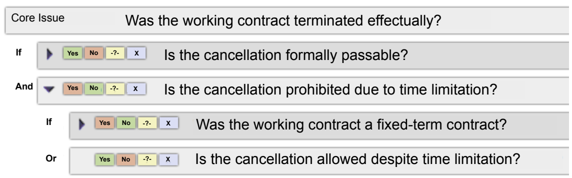

Figure 2: Exemplary *iTree* Implementation.

a diverse user population—i.e., users with different background and expertise might be able to profit from the same system. This is achieved by the possibility to derive the solution rating both by answering more abstract top-level questions (domain specialist level) or by stepping into more refined, elaborate questions (less expertise required) in *iTree*. By the visual representation of the knowledge base structure, moreover a form of focus-and-context view is created: Not only the currently active/processed question(s) are visible, but also surrounding elements are indicated—limited only by the display size. As the user thus can visually trace the result of an answer by the distinct presentation of the questions and their current state, that is propagated all throughout the tree, the core issue rating becomes more transparent. The chosen visual representation of the knowledge further supports users in

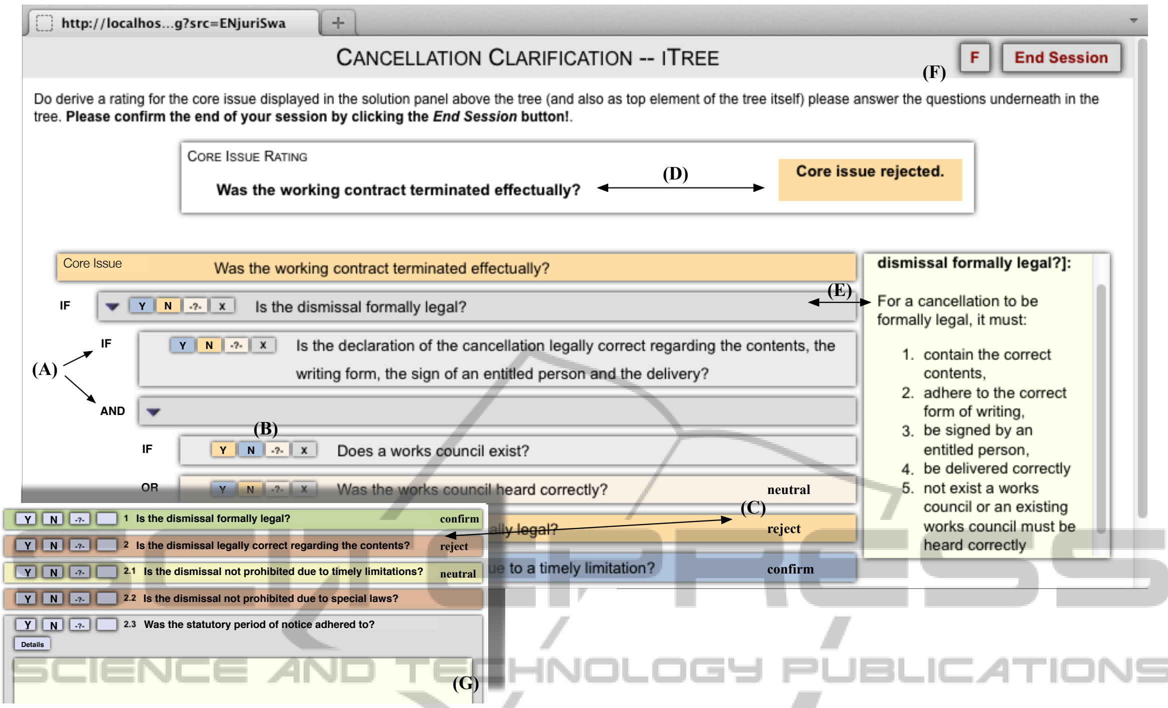

Figure 3: JuriSearch clarification module as *iTree* (large) and *oneQ* UI (small). AND/OR rules for rating the (sub-)questions (A) are visually represented; reversed question example (underneath B); *dummy* node example, only serving for rule connection modeling (above B); four simple buttons (B) for rating the questions (Y:yes—N:no—?:neutral—X/empty:retract), rating highlighted by background color (C); core issue rating prominently displayed and updated continuosly (D); additional information displayed in separate panel when mouse-overing question (E); anytime feedback/data collection features integrated in UI header (F); clarification core component in *oneQ* style (G) always displays current active question with additional-information panel, previously answered questions remain presented in a more condensed view.

gaining a thorough understanding of the investigated core issue and the coherences between its clarifying questions and the core issue itself. Thus, users acquire additional knowledge by means of the system, yet are also enabled to bring in their existing knowledge for potentially shortening the clarification session or for focussing on only those parts in detail that are rather unclear. Together with optional, auxiliary information that can be integrated for each of the elements (not contained in the scheme—see e.g. the auxiliary information panel in Figure 3, E), *iTree* specifically can serve as a skill-building UI type.

4 CASE STUDY—JuriSearch

At the beginning of 2012, the *JuriSearch* project was initiated as a cooperation between the university of Würzburg and the RenoStar corporation (partly founded by the Free State of Bavaria). JuriSearch aims at building a wKBS for the legal domain: The target system is intended to integrate both a standard consultation (entrance) module—*hypothesize*—and various clarification modules for each potential core issue—*test*—as to provide encompassing advice on various legal topics, such as the right of cancelation or the law of tenantry. Potential target users are diverse, ranging from legal laymen—searching for a basic understanding/estimation of their case to (fresh) lawyers seeking for guidance regarding legal (sub)domain(s) that are not exactly their special field of work. Those framing conditions provided a perfect opportunity to implement and evaluate the *iTree* UI style. Therefore, a comparative study with *iTree* and a more common, conversational UI style was conducted; the latter was implemented as a one-question UI style (*oneQ*). In contrast to the free, explorative interaction with *iTree*, *oneQ* is based on the metaphor of a conversation: The system always presents only the one suitable next question at a time, thus imitating a strict dialogue between a user and the system. Yet, both UI types are based on the exact same knowledge base with their core difference being the presentation of the questions: Hierarchical tree (*iTree*) vs. single question (*oneQ*). Refinement questions are as well available in *oneQ*; yet there, the former current question is folded and the first of the refinement level questions is presented, thereby destroying much of the 'contextual knowledge' that *iTree* facilitates by always presenting all questions of the current hierarchy level in addition to the surrounding, further structure (limited by display size only).

Figure 3 presents the *iTree* implementation in JuriSearch (A-F) as well as the alternative, conversational *oneQ* UI (G).

4.1 User Study—Framing Conditions

21 members from our department—all male, mostly between 25 and 35 years—participated in the first study; as computer scientists, they all were versed in general computer and web system usage, yet in most cases had little to no experience regarding the specific wKBS types, and no experience regarding the target domain at all. Two exemplary problem descriptions from the domain of cancellation were created, and participants were asked to solve one problem with iTree and the other with oneQ; to avoid biased results due to the sequence of using the UIs, that sequence was altered between participants. The study was conducted remotely: The test systems were deployed on a specifically configured server—enabling the integrated logging and feedback/questionnaire mechanisms—and the participants were given all required instruction material per email.

4.2 User Study—Results & Discussion

The collected log-data revealed a general applicability of the *iTree* concept for implementing a clarification wKBS UI in the legal domain and specifically the following results: First, *iTree* exhibited an average *task time* of 13m 38s±6m 49s in contrast to 10m 39s±5m 49s for *oneQ* (by a narrow margin statistically not significant on a one-sided unpaired t-test, p=0,068). The higher task time of *iTree* could possibly be explained by its ability to provide intuitively for free, extensive exploration of the system. Yet on the other hand, task time should not be overrated at all, here; the extent of usage of the test systems depended in larger parts on a) the reading speed of the participants regarding the questions and explanations, b) the usage conditions (during daily job routine vs. after end of work) which, due to the setting of the remote study could not be controlled strictly, and c) the potentially already existing knowledge regarding the problem at hand, in turn leading to highly subjective task time results between users. Regarding the *success-* and *error rate*, a case was classified successful, if the correct rating of the core issue was derived by the user with the respective system type, and not successful if either the wrong or no solution was found. For *iTree*, the success rate was 42,86% and 38,1% for *oneQ* (both: no statistical significance on a one-sided binomial test with p=0,11 and p=0,16); along with subjective user feedback, this clearly indicates the need to rework the knowledge base contents/structure for yielding better results. Furthermore, both anytime- and questionnaire-based feedback were collected as qualitative user data. The first remarkable finding was the fact, that *iTree* nearly concordantly was perceived more intuitively usable, and that it thus further was reported to be the preferred UI type by 81% of the study participants, whereas *oneQ* only was preferred by 14% and no preference was stated by 5%; this is statistically significant on a χ^2 test with $p<0{,}05$ and with an anticipated distribution of 50% (*iTree*), 30% (*oneQ*), and 20% (both equally). One possible explanation might be the specific characteristics of the participant population, that—as computer scientists—might simply be used to tree representations and thus perceived *iTree* as naturally more intuitive to use. Regarding further subjective (questionnaire) topics *iTree* scored better all over; on a scale from 0 (worst) to 6 (best) the results were: Comprehensibility of the system reactions 4.43±1.54 (*iTree*) vs. 2.76±1.45 (*oneQ*) or of the derived results 4.53±1.54 (*iTree*) vs. 3.33±1.85 (*oneQ*), and the mediation of domain knowledge to the user 4.05±1.32 (*iTree*) vs. 2.95±1.72 (*oneQ*); those differences are all statistically significant using an unpaired, one-sided t-test with $p\leq 0{,}05$. Especially the latter value affirmed our assumption that *iTree* particularly evinces skill-building abilities. Additional insights from anytime feedback included: The *wording of the questions* was perceived as incomprehensible/cumbersome in 11 cases (52%) due to often used duplicate negations and legal specialist language, probably further aggravated by the fact that the chosen participants were legal laymen and thus not at all familiar with legal terms and language; also, the hierarchical structure and representation of the knowledge base—that followed the legal subsumption logic—was perceived unfavorable. In such a hierarchy/sequence, the questions most interesting for legal laymen appear far down while at the same time more abstract concepts are contained at the upper levels; this led to (laymen) users having difficulties to make sense of the concepts at the top/beginning of the hierarchy/questioning sequence. A solution to this issue might be a complete restructuring of the knowledge base, so that the most relevant questions and distinctions—from the users' point of view, e.g., typical reasons for dismissal, size of company, etc.—also appear on rather top levels, sure posing a difficult trade off between legal correctness/-schematic thinking and understandability; yet it apparently could greatly contribute to tailoring the UI to the users in enabling them to bring in their own perspective and knowledge in the dialog. Thus a further refinement of the knowledge base with regards to a clear, easily understandable language and structure turned out indispensable. Another interesting finding was the fact, that in 4 (19%) cases, the real meaning of the *-?-* button as an answer alternative was not grasped; users rather expected the system to display more elaborate explanations on the issue at hand or to open up the next refinement level of the questions instead of receiving just a rating of the current question. Similarly, the *X/empty* button—designated to clearing a previously entered answer—was not intuitively understood in 3 (14%) cases.

5 CONCLUSIONS

In this paper, we claimed the importance of a careful UI/interaction design for web-based, knowledge-based systems. Regarding the consultation systems' sub-class clarification systems, we suggested *iTree* as novel UI/interaction style for increased efficiency and usability. In a first comparative user study from the legal domain, an initial *iTree* prototype as well as an alternative, one-question style prototype were implemented using the prototyping and knowledge systems engineering tool *ProKEt*. The results suggest, that *iTree* generally is a favorable UI style for clarification systems, that supports free, explorative system usage and thus provides skill-building potential on the side of the users. Yet, the study also showed the need to rework the knowledge base of the system, regarding both the question wording as well as their structuring. One assumption requiring further studies is, that the legal *iTree* at its current state is satisfying for legal experts, whereas a restructured system could be more appropriate for non-expert users. Additionally, we plan on developing and evaluating similar *iTree* systems for the medical domain. This raises the requirement of more fine-granular rating options, e.g., by scoring rules. Finally, further experimentation with potential UI enhancements is intended to help improve the *iTree* concept; one such idea is the integration of an interactive system state preview that is overlaid when mouse-overing the respective answer option.

ACKNOWLEDGEMENTS

We thank the RenoStar (Großwallstadt, GER) corporation for valuable discussions and cooperation.

REFERENCES

Angele, J., Fensel, D., Landes, D., and Studer, R. (1998). Developing Knowledge-Based Systems with MIKE. *Automated Software Engineering: An International Journal*, 5(4):389–418.

Baumeister, J., Reutelshoefer, J., and Puppe, F. (2011). KnowWE: A Semantic Wiki for Knowledge Engineering. *Applied Intelligence*, 35(3):323–344.

Brooke, J. (1996). SUS: A quick and dirty usability scale. In Jordan, P. W., Weerdmeester, B., Thomas, A., and Mclelland, I. L., editors, *Usability evaluation in industry*. Taylor and Francis, London.

Constantine, L. L. and Lockwood, L. A. D. (1999). *Software for Use: A Practical Guide to the Models and Methods of Usage-Centered Design*. Addison-Wesley Professional.

Duan, Y., Edwards, J. S., and Xu, M. X. (2005). Web-based expert systems: benefits and challenges. *Information & Management*, 42:799–811.

Freiberg, M. and Puppe, F. (2012). Prototyping-based Usability-oriented Engineering of Knowledge-based Systems. In *Proceedings of Mensch und Computer 2012 (to appear)*.

Freiberg, M., Striffler, A., and Puppe, F. (2012). Extensible prototyping for pragmatic engineering of knowledge-based systems. *Expert Systems with Applications*, 39(11):10177 – 10190.

Hart, S. G. (2006). Nasa-Task Load Index (Nasa-TLX); 20 Years Later. In *Human Factors and Ergonomics Society Annual Meeting*, volume 50.

Leichtenstern, K. and André, E. (2010). MoPeDT: features and evaluation of a user-centred prototyping tool. In *Proceedings of the 2nd ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, EICS '10, pages 93–102, New York, NY, USA. ACM.

Patil, S. S., Dhandra, B. V., Angadi, U. B., Shankar, A. G., and Joshi, N. (2009). Web based expert system for diagnosis of micro nutrients' deficiencies in crops. In *Proceedings of the World Congress on Engineering and Computer Science*, volume 1, pages 20–22.

Rahimi, S., Gandy, L., and Mogharreban, N. (2007). A web-based high-performance multicriteria decision support system for medical diagnosis: Research articles. *International Journal of Intelligent Systems*, 22:1083–1099.

Schreiber, G., Akkermans, H., Anjewierden, A., de Hoog, R., Shadbolt, N., de Velde, W. V., and Wielinga, B. (2001). *Knowledge Engineering and Management - The CommonKADS Methodology*. MIT Press, 2 edition.

Tomic, B., Jovanovic, J., and Devedzic, V. (2006). JavaDON: an open-source expert system shell. *Expert Systems with Applications*, 31(3):595 – 606.

URL d3web (2012). http://d3web.sourceforge.net/ (last checked Apr. 15th, 2012).