# MAUSE

## Towards the **MA**turation of Information Technology **US**ability **E**valuation

**http://cost294.org/**

## What is COST294-MAUSE?

COST294-MAUSE is a usability research community that is run under the auspices of COST. Each COST Action is assigned a unique number and ours is 294. MAUSE stands for Towards the MAturation of IT USability Evaluation – the overarching theme of the community. The ultimate goal of COST294-MAUSE is to bring more science to bear on Usability Evaluation Methods (UEMs) development, evaluation, and comparison, aiming for results that can be transferred to industry and educators, thus leading to increased competitiveness of European industry and benefit to the public. We aim to achieve three main objectives:

• To deepen the understanding of the inherent strengths and weaknesses of individual UEMs;

• To identify reliable and valid methods to compare different UEMs in terms of their effectiveness, efficiency as well as scope of applicability;

• To develop efficacious strategies for extracting useful information from the results of UEMs to improve the system test.

The major rationale for our collaboration is that we share the vision to improve the research as well as practical work on usability. Existing problems in usability research, from basic to intricate, need to be resolved through extensive co-operation within a community of usability professionals and researchers with diversified backgrounds. In fact, usability research has been rather fragmented and scattered in a variety of industrial and academic institutions. It is high time to coordinate these distributed efforts to best utilize the resources available and to mutually stimulate and enrich ongoing research activities. COST294-MAUSE is a community committed to address this challenge.

## Members and Expertise

Currently, there are 37 management committee and 36 additional members from 35 different European organizations. Most of the members are experienced researchers in HCI and some are postgraduate students pursuing their studies in the area of usability.

As diverse as the issues in usability, are areas of expertise possessed by the MAUSE members. So are the research topics they are working on. The range of application domains where they investigate their topics is quite wide: safety critical systems such as aviation traffic control and nuclear power process control, digital libraries and e-learning, e-commerce, assistive technologies such as universal accessible websites and haptic devices, mobile HCI, medicine and film-making.

The various backgrounds and divergent foci of the members' works are complementary and beneficial to the fulfillment of the objectives of the MAUSE community.

MAUSE members have established a relatively extensive network of liaisons with other research bodies in the academic and industry within and without Europe. They are also active members of relevant professional associations such as UPA, ACM SIGCHI, IFIP TC13 Working Groups 13.1/13.2/13.4/13.5 etc., and have successful collaboration with a number of private enterprises such as telecoms companies and banks.[2]

## Organizational Structure

The technical work of COST294-MAUSE is driven by Chairperson, Vice-Chairperson, Secretary, Scientific-Coordinator, Dissemination- and Working Group-Coordinators and is supported by all the partners from the signatory countries. The Action was officially founded in January 2005 and its activities will be funded by COST until January 2009. As of April 2006, there are 20 signatory countries with each of them having at most two representatives in the Management Committee (MC).



COST294 Management Committee
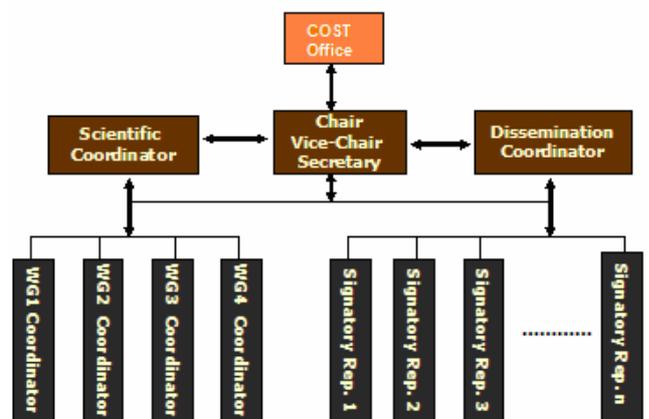
## Tools and how to get informed

**MAUSE Digital Library:** http://cost294.org/dl.php
The public part contains both draft and final version of working documents as well as deliverables of the project. Recently added documents include instruments for evaluating and slides illustrating the interconnections between the four WGs and the interfaces between the WGs and partner projects. The private part (restricted to MAUSE members) contains draft documents that are currently under development by the working groups.

**MAUSE Mailling List:**
http://listes.irit.fr/wws/info/cost294-announcement
To keep informed about news, events and current activities, join MAUSE cost294-announcement mailing list.

**MAUSE Newsletter:** http://cost294.org/newsletter/
Contains the electronic version of all newsletter.

http://www.esf.org/

http://cost.cordis.lu/

## WG1: Critical Review and Analysis of Individual UEMs

**Coordinator: Dominique Scapin**, **INRIA**, **France**

WG1 has developed two templates, namely *Usability Methods (UM) Generic Description* and *Case Study Description*, which have been undergone several iterations of improvements. Put briefly, the *UM Generic Description* template is to critically analyze a method with a set of carefully selected attributes that enable thorough understanding of the usability method of interest. For instance, a UM is first classified as one of three major categories: DGMM (Data Gathering & Modelling Methods), UIEM (User Interactions Evaluation Methods) or CMs (Collaborative Methods). Each of these categories is further divided into finer sub-categories. Then the UM is reviewed meticulously in terms of their values for an array of attributes such as theoretical background, advantages, drawback, estimated impact, to name just a few. The «Case Study Description» is to describe the conditions and contexts under which the method is implemented. Thus far eight UMs have been analyzed. An interim technical report based on the above analyses has been published and accessible via the MAUSE Digital Library. The WG1 work will be extended by:

• Enlarging the coverage of a variety of UEMs

• Inviting more submissions on case studies

• An easier paper-like word version is now available

• Consolidating results as input to other three WGs

• Identifying effective means to deliver the knowledge and experience gained



**First MC+WG Meeting at the University of Sunderland, UK**

## WG2: Comparing UEMs: Strategies and Implementation

**Coordinator: Gilbert Cockton**, **University of Sunderland**, **UK**

WG2 aims to tackle three major tasks with respect to its overarching goal of comparing usability evaluation methods:

**Task 1**: to develop comprehensive, consistent and robust means to compute measures from a range of constructs for use when comparing methods to each other, but with no measures of thoroughness or validity. Constructs and associated methods for coding and rating will be the key task foci.

**Task 2**: develop quality models for the Web and compare their performance to UEMs;

**Task 3**: to develop consistent and robust means to compute thoroughness and validity relative to a comprehensive, rigorously derived and well grounded set of problems that have been actually observed and detected in real use (i.e., a focus on the problem of accurate coding of false positives and creating a maximal ("Being the greatest or highest possible") actual problem set for thoroughness computations. Problem extraction and merging, and maximization of actual problem sets will be the key task foci;

For Task 1, 18 comparison constructs have been proposed, including stages of interaction, severity, frequency, cognitive resources for discovery and analysis, problem report format, downstream utility, to name just a new. For each construct, definition, review of usage as well as of associated coding and rating approaches will be undertaken. The inter-dependencies among the constructs will also be studied.

For Task 3, review and solution proposal will be undertaken for methodological issues associated with collecting robust data from which thoroughness and validity can be derived and for specific issues associated with problem merging, structuring and comparison. Furthermore, the concept of reliability will be extended beyond effectiveness as a simple function of thoroughness and validity.

For Task 2, software quality models will be reviewed to identify potential constructs for usability evaluation method assessment and comparison.

Furthermore, multi-site experiments will be coordinated to address Task 1 and Task 3. In particular, two such experiments on usability problem merging and on coding and rating of comparison constructs are under preparation. WG2 will have ongoing collaboration with the three WGs.

## WG3: Refining and Validating Classification Schemes for Usability Problems

**Coordinator: Mark Springett**, **Middlesex University, UK**

The primary objective of this Activity is to improve and substantiate the tools for analysing usability problems.

The existing defect classification schemes will be reviewed to evaluate their suitability for isolating causes and proposing redesign solutions. A list of possible but not exhaustive schemes includes: Orthogonal Classification Scheme (ODC), Root Cause Analysis (RCA), Hewlett Packard Defect Classification Scheme (HP-DCS), Usability Problem Taxonomy (UPT), User Action Framework (UAF), Classification of Usability Problem Scheme (CUP).

Initial validations of the CUP DCS handbook have been carried out and a second version of the handbook produced. The design for a major multi-site validation experiment is complete and volunteer participants covering a range of application domains have been recruited. The work will also include a comparison with other classification schemas such as the User-action Framework. This work is scheduled to be complete by the end of 2006. Analysis of Quality Factors: This thread of work intends to integrate analysis of user-experience into the specification of non-functional requirement and the setting of evaluation targets. Initial work on this thread consisted of a literature search in which an agenda of theoretical issues was established, as well as investigation of candidate practical techniques suitable for elicitation of and evaluation against key quality attributes. A template is now prepared detailing a multi-site experiment in which the Repertory Grid technique will be used to elicit key constructs expressed by domain professionals, design team members and end-users. Volunteer organisations are currently being recruited to partake in the study, which is expected to be complete in early 2007.

The study will also incorporate a specific sub-theme in which perceptions and interpretations of accessibility are probed. The long-term aim of this work is to recommend approaches for integrating accessibility with design practice in a way that appropriately represents the requirements across the spectrum of users. A separate report on this study will be published in 2007.

## WG4: Review on Computational and Definitional Approaches in Usability Evaluation

**Coordinator: Christian Stary, University of Linz, Austria**

Working Group 4 (WG4) is concerned with a systematic review on the computational and definitional approaches in usability evaluation. Our primary target is a systematic review of existing models and procedures for estimating or calculating certain key usability test parameters and traditionally defined usability quality metrics. Of primary interest are the capabilities of formal representations and corresponding computational schemes. Computational approaches in usability evaluation target towards the automated evaluation of user interfaces, user characteristics, scenarios of use, or specifications. Since the proposed variety of techniques show different backgrounds and various conceptual and operational assumptions, it is difficult to find out their specific capabilities which determine their quality. To facilitate the review process a layered scheme for the structured evaluation of existing computational usability evaluation methods (UEMs) has been proposed. In contrast to existing schemes it reveals relationships between usability principles and computable measurements. The scheme can be seen as an implementation of some comparative analysis principles (CAP) which are shortly described in the following: CAP1: Computational UEM Relevance, CAP2: Definitional Primacy for Usability Evaluation, CAP3: CUEM Context Specification, CAP4: Scope of CUEM, CAP5: Explication of Semantic Encoding(s), CAP6: CUEM Accuracy Determination.

WG 4 is currently finishing their review process. More than ten CUEMs are expected to be analyzed. The CUEMs has been chosen by checking whether direct, or at least interpreted statements to the different categories of our review scheme - which is an implementation of the above described principles - can be found in the published literature. The reviewing results indicate so far, that a wide range of different methods are used within the different CUEMs. In the next step we will compile a meta analyses which is based on the single CUEM reviews. Therefore the reviewing results, that are documented in our review-scheme has to be interpreted in terms of the above CAPs. The result will be documented in a meta scheme and work as the basis for defining the maturation of computational usability evaluation.

### Related Conferences

**ACM CHI: Conference on Human Factors in Computing Systems**
**Location:** Montréal, Québec, Canada
**Dates:** April 22nd – 27th 2006
**Web site:** http://www.chi2006.org/

**AVI: Advanced Visual Interfaces**
**Location:** Venice, Italy
**Dates:** May 23rd - 26th 2006
**Web site:** http://www.dsi.unive.it/avi2006/

**Usability Professionals 06**
(Organized by the German Chapter of the Usability Professionals' Association in compüoperation with "Mensch und Computer")
**Location:** Gelsenkirchen, Germany
**Dates:** September 3rd – 6th 2006
**Web site:** http://www.usabilityprofessionals.de/

**ROCHI: The 3rd National Conference on Human-Computer Interaction 2006** (annual ACM SIGCHI Romania forum)
**Location:** Bucharest, at the Academy of Economic Studies
**Dates:** September 21st-22nd 2006
**Web site:** http://rochi.utcluj.ro/

**NordiCHI: 4th Nordic Conference on Human-Computer Interaction**
**Location:** Oslo, Norway
**Dates:** October 14th – 18th 2006
**Web site:** http://nordichi.net.dynamicweb.dk/

**Regina Bernhaupt**
*ICT&S-Center, Universität Salzburg*
*Salzburg, Austria*

*Formal Methods and User Testing? Antithesis or Synergy?*

Can formal methods support traditional user testing? What are the connection points between these two branches of HCI? To clear these central questions Regina Bernhaupt spent four weeks during February 2006 at IRIT, University Paul Sabatier, Toulouse, France.

The LIIHS research group (lead by Philippe Palanque with their members Remi Bastide, David Navarre, Marco Winckler, …) is currently working on formal models for multimodal interfaces. Regina's background is in traditional user testing for all kinds of interaction techniques and ethnographic studies. The goal of the stay was to develop a new usability evaluation method in the intersection of formal methods and methods for evaluating user experiences.

The results of this stay are a book chapter on combining formal methods and traditional methods for evaluating multimodal interfaces (MMI) and a publication accepted at the SpaceOps Conference 2006. Another outcome is the participation of the University of Salzburg, ICT&S Center, HCI & Usability Unit in the project IMAGE (development of multimodal interfaces) in 2006, to test the newly developed method.

The visit in Toulouse was a good opportunity to connect two branches of HCI normally not working closely together. I hope that this cooperation will lead to the development of new evaluation methods suitable not only for multimodal interfaces. Thanks to the team for the nice support, for their efforts in communicating in English with me and all their help to make it a pleasant stay for me.

**Svetlena Taneva**
*ETH Zürich*
Zurich, Switzerland

*How can medical systems design learn from Air-Traffic control systems design?*

This was the question which lead me to LIIHS at University Toulouse 3 in France. I learned about the design and usability evaluation methodologies being developed at LIIHS for Air Traffic Control systems, and safety-critical systems in general. Then, together with the local team, we explored the idea of transferring and modifying these methodologies to tailor them to the medical domain. The key area we focused on was model-driven software engineering that accounts for potential human error and communication breakdown. Specifically, we looked at how one can incorporate knowledge about errors into the design of a system in order to produce a more error-tolerant and safer medical application.

We applied our ideas to a case study - the adverse events reports associated with a telemetry patient monitoring system made by a major medical manufacturer. We showed how our approach could have predicted and prevented numerous adverse events. Over the last thirty months, 19 out of 21 reports about this system submitted to the U.S. Food and Drug Administration (FDA) Manufacturer and User Facility Device Experience database (MAUDE) are associated with communication breakdowns; 89% of the breakdowns resulted in a life-threatening situation or in the patient's death. This could have been avoided. An adverse event minded approach to design would have contributed to a safer and more error tolerant telemetry system.

Our cooperation continues beyond my short stay in Toulouse – we are currently writing a paper with our results. Overall, it was a great learning experience and a very productive visit. My thanks to MAUSE for making this possible!

## Collaboration Instruments in MAUSE

Meetings of Management Committee (MC) are held in conjunction with those of the four WGs on a regular basis. The primary goals of the meetings are to deal with emerging issues and needs, to consolidate the ideas and findings that have been collected in the meantime, and to identify new research areas that are necessary and worthy to pursue.

To deepen the members' understanding of specific topics, experts have been invited to present a talk in the meetings, including:

• Michael Harrison (University of New Castle upon Tyne, UK) on "The Pros and Cons of Mathematical Models in Interactive Systems"

• Annelise Mark Pejtersen (Risø National Laboratory, Denmark) on "Cognitive Task Analysis"

• Marc Hassenzahl (Technical University of Darmstadt, Germany) on "User Experience and Hedonic Quality".

Apart from regular MC + WG meetings, the other instruments of collaboration are deployed:

Annual Open Workshop: It is a valuable instrument to facilitate a link between the MAUSE community and the HCI community at large. In anticipation of international attendance, our workshops are normally held in conjunction with established HCI conferences and calls for participation are issued to different channels.

Each workshop focuses on a specific theme:

Our first workshop, with the theme '**User Interface Quality Models**' **(UIQM)**, was held in conjunction with INTERACT'05. 34 participants distributed into the four groups discussed enthusiastically the issues concerned, thereby deepening their understanding as well as stimulating further thoughts pertaining to the theme of the workshop. For further information see: http://cost294.org/uiqm2005/

**Short Term Scientific Mission (STSM)**: STSMs are exchange visits of researchers to the institutes of the community members to perform some specific studies or experiments, thereby enabling mutual enhancement of expertise and experience. Two STSMs have successfully been implemented (see STMS at page 3).

**Multi-site Experiments (MSE)**: MSE are coordinated to validate research ideas and concepts on a large-scale across expertise, domains and cultural contexts. Several MSEs on different themes are currently being organized, including validation studies on merging, coding and rating problem sets as well as usability problems classification.

**Training/Summer School**: It offers Ph.D. students the opportunity to know more the current development of the field. Currently, the opportunities to realize this instrument are being explored.

### Partners

**Austria**: University of Linz, ICT&S/ Universität Salzburg
**Belgium**: Catholic University of Louvain, SmalS
**Cyprus**: University of Cyprus
**Denmark**: University of Copenhagen, University of Southern Denmark- Odense
**Finland**: University of Oulu, Technical Research Centre of Finland (VTT)
**France**: INRIA-Rocquencourt, LIIHS-IRIT, University of Paris-10
**Germany**: Augsburg University of Applied Science, University of Konstanz, University of Technology Darmstadt
**Greece**: Aristotle University, University of Piraeus
**Iceland**: University of Iceland, Reykjavik University
**Italy**: Universita' di Bari, ISTI-CNR
**Poland**: Gdansk University of Technology
**Romania**: ICI Bucuresti, Babes-Bolyai University
**Norway**: SINTEF
**Serbia & Montenegro**: University of Novi Sad
**Spain**: Technical University of Valencia, Universidad de Castilla-La Mancha
**Slovenia**: Institut Jozef Stefan
**Sweden**: Uppsala University
**Switzerland**: ETH Zürich, Inst. für Software-Ergonomie und Usability, Medizinische Fakultät der Universität Bern
**The Netherlands**: TU Delft
**United Kingdom**: Middlesex University, Sunderland University, University College London Interaction Centre
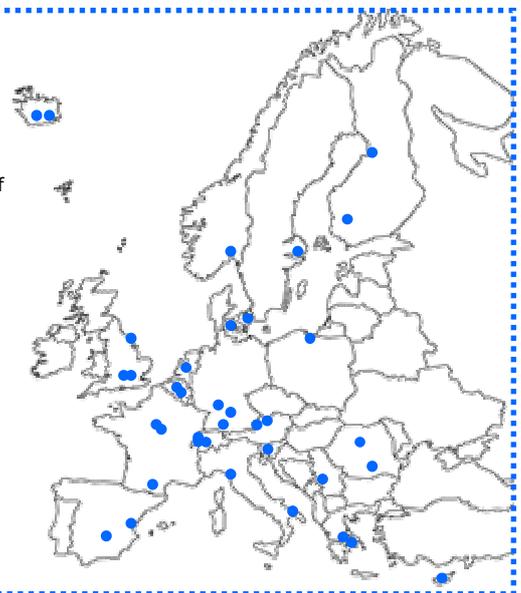
## MAUSE Book: "Maturing Usability: Quality in Software, Interaction and Values"

Our first book project entitled "**Maturing Usability: Quality in Software, Interaction and Values**" is an edited collection of contributions authored by leading HCI researchers and practitioners.

The book will consist of four major parts with the first three addressing specific ideas and concerns germane to the three respective scopes reflected in the title and the fourth part reviewing and projecting the future of usability research and practice.

The international authorship, the editorial board and the publisher (Springer Verlag) will closely collaborate to produce a highly useful volume. It is scheduled for publication in early 2007.

### Upcoming MAUSE Events

**Second Annual Open Workshop:** A proposal entitled "*User Experience – Towards a unified view*" has been submitted. Upon acceptance, it will be held in conjunction with the upcoming **NordiCHI'06 14-18 October 2006** in Norway. User Experience (UX) is an emerging area with a number of theoretical, methodological and empirical issues entailing more research efforts. Of particular concerns are how to integrate the existing evaluation methods and conceptual models pertaining to UX.

**Fourth and Fifth MC + WG Meetings:** The two meetings are tentatively schedule to be held in collocation with NordiCHI'06 in Norway and in Sweden (Uppsala University), respectively.

### MC Chair & co-Chair

Effie Law [1] & Ebba Hvannberg [2]
[1] ETH Zürich, Institute TIK
[ law@tik.ee.ethz.ch ]
[2] University of Iceland
[ ebba@hi.is ]

### Dissemination Activity

Marco Winckler & Philippe Palanque
[ winckler, palanque ]@irit.fr
LIIHS-IRIT, Université Toulouse 3
118 Route de Narbonne
31062 Toulouse Cedex 4, France

### Other Events

**Opening of the usability and experimental lab, ICT&S Center**
Location: University of Salzburg
Dates: 3rd May 2006,
Web: http://www.icts.uni-salzburg.at/ > Events