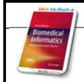


Andreas Holzinger
VO 709.049 Medical Informatics
20.01.2016 11:15-12:45

Lecture 12

Methodology for Information Systems: Usability and Evaluation

a.holzinger@tugraz.at
Tutor: markus.plass@student.tugraz.at
http://hci-kdd.org/biomedical-informatics-big-data

Schedule

- 1. Intro: Computer Science meets Life Sciences, challenges, future directions
- 2. Back to the future: Fundamentals of Data, Information and Knowledge
- 3. Structured Data: Coding, Classification (ICD, SNOMED, MeSH, UMLS)
- 4. Biomedical Databases: Acquisition, Storage, Information Retrieval and Use
- 5. Semi structured and weakly structured data (structural homologies)
- 6. Multimedia Data Mining and Knowledge Discovery
- 7. Knowledge and Decision: Cognitive Science & Human-Computer Interaction
- 8. Biomedical Decision Making: Reasoning and Decision Support
- 9. Intelligent Information Visualization and Visual Analytics
- 10. Biomedical Information Systems and Medical Knowledge Management
- 11. Biomedical Data: Privacy, Safety and Security
- **12. Methodology for Info Systems: System Design, Usability & Evaluation**

Learning Goals: At the end of this 12th (last) lecture you …

- … understand the concepts and importance of usability
- are aware that medical software is now included within the Medical Device Act (Medizinprodukte-Gesetz, MPG);
- have a feeling for quality and can determine between product quality, process quality and information quality;
- are familiar with some important ISO standards for quality and usability of medical software and systems;
- understand the user-centered design process, from concept phase till verification and validation;
- are able to apply some usability engineering methods and evaluation methods applicable in the medical domain;
- understand the importance of evaluation and benchmarking (cost – time – quality), & again the ROC ☺

Keywords of the 12th Lecture

- Action analysis/Cognitive walkthrough
- Emotion recognition
- Ergonomics
- Hedonomics
- **Evaluation/Benchmarking: Accuracy, Precision, Validity, Reliability**
- Human-Centered Design (HCD)
- Medical Device Directive (MDD)
- Medical Product Law
- Medical Software
- Medizin Podukte Gesetz (MPG)
- **Quality**
- Software quality
- Technology Acceptance Model (TAM)
- Thinking aloud
- Usability Engineering (UE)
- User-Centred Design (UCD)
- Validation
- Verification

Advance Organizer (1/4)

- **Accessibility** = the degree to which a system or service is available to a diverse set of end users;
- **Accreditation** = a formal declaration by the Accreditation Authority that a system is approved to operate in the defined standards with accuracy, completeness and traceability;
- **Act** = a formal law passed by a legislative body;
- **Audit** = is performed to verify conformance to standards by review of objective evidence (e.g. ISO 9001), it is an independent examination of the life cycle processes within the audited organization;
- **Certification** = a (product/software) qualification to verify that performance tests and quality assurance tests or qualification requirements are certified;
- **cognitive modeling** = aka mental modeling = producing a computational model for how people perform tasks and solve problems, based on psychological principles. These models may be outlines of tasks written on paper or computer programs which enable us to predict the time it takes for people;
- **cognitive walkthrough** = an approach to evaluating a user interface based on stepping through common tasks that a user would need to perform and evaluating the user's ability to perform each step;
- **Consistency** = principle that things that are related should be presented in a similar way and things that are not related should be made distinctive.
- **consistency inspection** = a quality control technique for evaluating and improving a user interface. The interface is methodically reviewed for consistency in design, both within a screen and between screens, in graphics (color, typography, layout, icons), text (tone, style, spelling);
- **Effectiveness** = the degree to which a system facilitates a user in accomplishing a specific task, measured by task completion rate; often confused with efficiency;
- **Efficiency** = a measurable concept, determined by the ratio of output to input; it is the ability to accomplish a task in minimum time with a minimum of effort (once the end users have learned to use the system); often confused with effectiveness;
- **Emotion** = a mental and physiological state associated with a wide variety of feelings, thoughts, and behaviors, very important for usability;
- **end user** = the primary target user of a system, assumed to be the least computer-literate user;

Advance Organizer (2/4)

- **End-user programming (EUP)** = making computational power fully accessible to expert end users, e.g. to medical professionals with no specific computer programming knowledge; usually done by a user interface which enables easy programming (e.g. visual programming, natural-language syntax, wizard-based programming, mash-up programming);
- **Errors** = an important measurement of usability on how many errors do end-users make, how severe are these errors, and how easily they can recover from the errors;
- **Evaluation** = is the systematic process of measuring criteria against a set of standards;
- **Formative Evaluation** = usability evaluation that helps to "form" the design process, i.e. evaluation is taking place parallel and iteratively to the development process;
- **Heuristic Evaluation** = method to identify any problems associated with the design of user interfaces;
- **ISO 13407** = Human Centred Design Processes for Interactive Systems;
- **ISO 13485 (2003)** = represents the requirements for a comprehensive management system for the design and manufacture of medical devices;
- **ISO 14971 (2007)** = risk management for medical devices;
- **ISO 62304 (2006)** = Medical device software;
- **ISO 9001** = The ISO 9000 international standards family is for quality management and guidelines as a basis for establishing effective and efficient quality management systems;
- **ISO 9241** = Software usability standard;
- **ISO 9241-10** Ergonomic requirements for office work with visual display terminals (VDTs): Dialogue principles (1996);
- **ISO 9241-11** Ergonomic requirements for office work with visual display terminals (VDTs): Guidance on usability specifications and measures (1998);
- **ISO/HL7** = joint ISO and HL7 (Health Level Seven) International Standard;
- **ISO/IEEE** = joint ISO and IEEE (Institute of Electrical and Electronics Engineers) International Standard;
- **ISO/OECD** = joint ISO and OECD (Organisation for Economic Cooperation and Development) International Standard;

Advance Organizer (3/4)

- **Learnability** = degree of which a user interface can be learned quickly and effectively by measure of learning time;
- **learning curve** = the amount of time an end-user needs to fulfill a previously unknown task;
- **Mash-up** = the use of existing functionalities to create new functionalities, Mash-up composition tools are usually simple enough to be used by end-users without programming skills (e.g. by supporting visual wiring of GUI widgets, services and/or components together); The concept of mash-up are combination, visualization and aggregation in order to make data useful;
- **Medical Safety Design** = process including usability engineering and risk management to make the product compliant to EN 60601 and EN 62366 which is no longer a nice to have, but a requirement; the developer must provide a documentation on the usability engineering process;
- **Medizin Produkte Gesetz (MPG)** - Medical device act = valid law in Austria, based on European law (in Germany: Medizinproduktegesetz MPG in der Fassung der Bekanntmachung vom 7. 8. 2002 (BGBl. I S. 3146), das durch Artikel 13 des Gesetzes vom 8. 11. 2011 (BGBl. I S. 2178) geändert worden ist);
- **Memorability** = the measure of when an end-user returns to the system after a period of not using it, how easily can he re-establish efficiency;
- **Mental model** = the internal model of an end user on how something works; can be used by the designer for aligning his design strategy with human behavior;
- **Methodology** = systematic study of methods that are, can be, or have been applied within a discipline;
- **Participatory design** = a common approach to design that encourages participation in the design process by a wide variety of stakeholders, such as: designers, developers, management, users, customers, salespeople, distributors, etc;
- **Performance** = measurement of output or behaviour in both engineering and computing;
- **Performance measure** = a quantitative rating on how someone performed a task, such as the time it took to complete, the number of errors they made in doing it, their success rate, time spent in a particular phase of a process;
- **Satisfaction** = a subjective degree of how much an end-user enjoys using a system (joy-of use, enjoyability);

Advance Organizer (4/4)

- **Semiotics** = the study of signs and symbols and their use in communicating meaning, especially useful in analyzing the use of icons in software, but also appropriate to the analysis of how screen design as a whole communicates;
- **Software Usability Measurement Inventory (SUMI)** = a rigorously tested and proven method of measuring software quality from the end user's point of view; consistent method for assessing the quality of use of a software product or prototype;
- **Software Usability Scale (SUS)** = a ten-item attitude Likert scale providing a single score reflecting the overall view of subjective assessments of usability, developed by Brooke (1986), the power is in its simplicity;
- **Task analysis** = a set of methods for decomposing people's tasks in order to understand the procedures better and to help provide computer support for those tasks;
- **Thinking aloud** = direct observation, where end-users are asked to speak out loud everything they do, think, feel in each moment during execution of a task; the only method to gain insight into the thinking, helpful at early stages of design for determining expectations and identifying what aspects of a system are confusing;
- **Usability engineering** = a methodical approach to user interface design and evaluation involving practical, systematic approaches to developing requirements, analyzing a usability problem, developing proposed solutions, and testing those solutions;
- **User Interface (UI), Graphical User Interface (GUI)** = input/output possibilities of a system - for the end-user, the interface actually *is* the system;
- **Validation** = is a (external) quality process to demonstrate (to the stakeholder) that the system complies with the original specifications;
- **Verification** = is a (internal) quality process, used to evaluate whether and to what extent the system complies with the original specifications;

Slide 12-1 Key Challenges

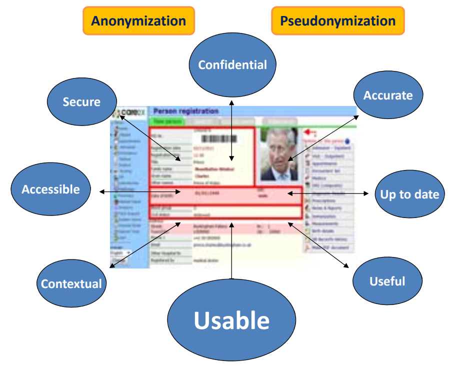
- **Usability, Accessibility, Reliability** are still underestimated in health applications [1]
- **User-Centred Designs** are rarely applied in medical information systems [2]
- **Evaluation and Benchmarking** are of utmost importance – but use statistical benchmarking with care! [3]

[1] Holzinger, A. 2005. Usability engineering methods for software developers. Communications of the ACM, 48, (1), 71-74.

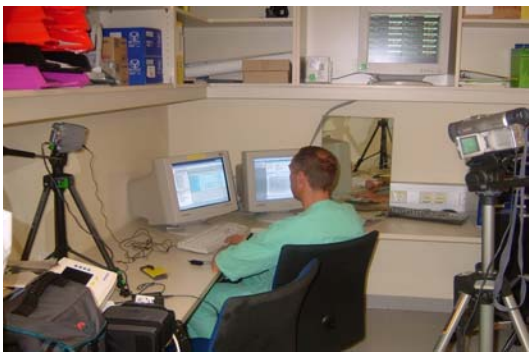
[2] Thimbleby, H. 2007. User-Centered Methods Are Insufficient for Safety Critical Systems. Lecture Notes in Computer Science (LNCS 4799). Springer, pp. 1-20.

[3] Drummond, C. & Japkowicz, N. 2010. Warning: statistical benchmarking is addictive. Kicking the habit in machine learning. Journal of Experimental & Theoretical Artificial Intelligence, 22, (1), 67-80.

Please remember:

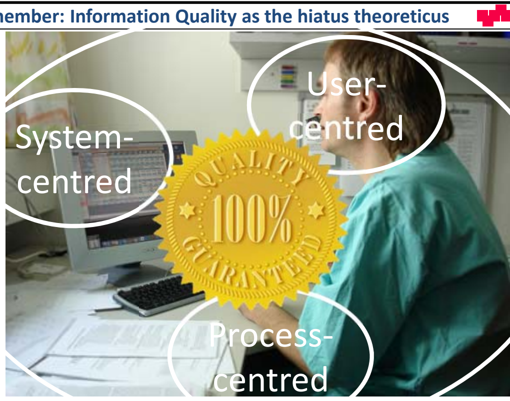


Slide 12-2: Medical Workplace Usability - enhance quality



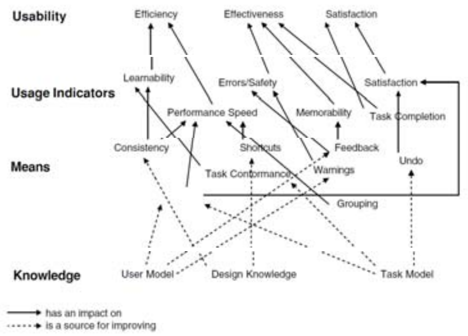
Holzinger, A. & Leitner, H. (2005) Lessons from Real-Life Usability Engineering in Hospital: From Software Usability to Total Workplace Usability. In: *Empowering Software Quality: How can Usability Engineering reach these goals? Vienna, Austrian Computer Society, 153-160.*

Remember: Information Quality as the hiatus theoreticus



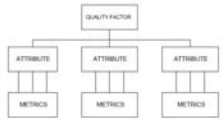
Holzinger, A. & Simonic, K.-M. (Eds.) (2011) *Information Quality in e-Health. Lecture Notes in Computer Science LNCS 7058, Heidelberg, New York, Springer.*

Slide 12-3: A framework for understanding usability



Veer, G. C. v. d. & Welie, M. v. (2004) DUTCH: Designing for Users and Tasks from Concepts to Handles. In: Diaper, D. & Stanton, N. (Eds.) *The Handbook of Task Analysis for Human-Computer Interaction. Mahwah (New Jersey), Lawrence Erlbaum, 155-173.*

Slide 12-4: System characteristic versus Quality factor



System Characteristic	Corresponding Quality factor(s)
Safety-critical (medical) Systems	Reliability, Correctness, Verifiability
Classified (patient) data	Security
Real-time operation	Efficiency
Heterogeneity of system landscape	Portability
Diverse set of (medical) end users	Usability
Possible further (hospital) development	Expandability

Cf. with: Cosgriff, P. (1994) Quality assurance of medical software. *Journal of Medical Engineering & Technology, 18, 1, 1-10.*

Slide 12-5: ISO Standards for Healthcare

ISO/TC 215 Health informatics

About | Contact details | Structure | Liaisons | Meetings | Tools

Secretariat: ANSI
Secretary: Ms Lisa Spellman
Chairperson: Mr Michael Glickman until end 2017
ISO Technical Programme Manager: Dr Mary Lou Pelaprat
ISO Editorial Programme Manager: Mrs. Laura Mathew
Creation date: 1998

Scope:

Standardization in the field of health informatics, to facilitate the coherent and consistent interchange and use of health-related data, information, and knowledge to support and enable all aspects of the health system.

Total number of published ISO standards related to the TC and its SCs (number includes updates):	152
Number of published ISO standards under the direct responsibility of ISO/TC 215 (number includes updates):	152
Participating countries:	31
Observing countries:	28

http://www.ahima.org/

An introductory video about ISO and healthcare: https://youtu.be/3-8nuqRo3-M

Slide 12-6: EU Directive 93/42/EEC Medical Device (MDD)

- The EU directive 93/42/EEC1 states criteria to define medical devices. For systems and devices that fall under these definitions, the directive states requirements that have to be met.
- Medical devices in the sense of the directive are devices that serve the following purposes:
- 1) Diagnosis, prevention, monitoring, treatment or alleviation of disease,
- 2) Diagnosis, monitoring, treatment, alleviation of or compensation for an injury or handicap,
- 3) Investigation, replacement or modification of the anatomy or of a physiological process,
- 4) control of conception;
- **The important aspect for IT systems is that software of medical devices is explicitly included in this definition.**
- Every device classified a medical device under the above criteria has to bear a CE 2 (conformité européenne) mark

Neuhaus, C., Polze, A. & Chowdhuryy, M. M. R. (2011) *Survey on healthcare IT systems: standards regulations and security (Technical report) Potsdam, Hasso-Plattner-Institute for Software Engineering.*

Slide 12-7: Quality of Med Software – standards to know

- ISO 9241 Software Usability
- ISO 13407 Human-Centred Development
- ISO 14971:2007 Risk Management
- Medical Device Act MPG (2010) incl. Software
- ISO 62304:2006 Medical Software
- UPA (2000) Life Cycle Processes
- ISO 27799:2008 Health informatics Information security management
- ISO 13485:2003 Medical Product Quality
- EU 93/42 Medical Device Directive (MDD)

12-8: MPG (Medizin Produkt Gesetz) includes Software …

BUNDESGESETZBLATT
FÜR DIE REPUBLIK ÖSTERREICH

Jahrgang 2009 Ausgegeben am 30. Dezember 2009 Teil I

143. Bundesgesetz: **Änderung des Medizinproduktegesetzes und des Arzneimittelgesetzes (NR: GP XXIV RV 466 AB 549 S. 49. BR: AB 8236 S. 780.) [CELEX-Nr.: 32007L0047, 32009L0120]**

143. Bundesgesetz, mit dem das Medizinproduktegesetz und das Arzneimittelgesetz geändert werden

Der Nationalrat hat beschlossen:

Artikel 1
Änderung des Medizinproduktegesetzes

Das Medizinproduktegesetz – MPG, BGBl. Nr. 657/1996, zuletzt geändert durch das Bundesgesetz BGBl. I Nr. 77/2008 und die Bundesministeriengesetz-Novelle 2009, BGBl. I Nr. 3, wird wie folgt geändert:

1. Im § 2 Abs. 1 lauten die Einleitungsworte:

„Medizinprodukte" sind alle einzeln oder miteinander verbunden verwendeten Instrumente, Apparate, Vorrichtungen, Software, Stoffe oder anderen Gegenstände, einschließlich der vom Hersteller speziell zur Anwendung für diagnostische oder therapeutische Zwecke bestimmten und für ein einwandfreies Funktionieren des Medizinprodukts eingesetzten Software, die vom Hersteller zur Anwendung für Menschen für folgende Zwecke bestimmt sind:"

Slide 12-9 Medical Product Law and mobile Apps

http://www.informationweek.com/desktop/medical-apps-on-tablets-gain-popularity

Quality first!

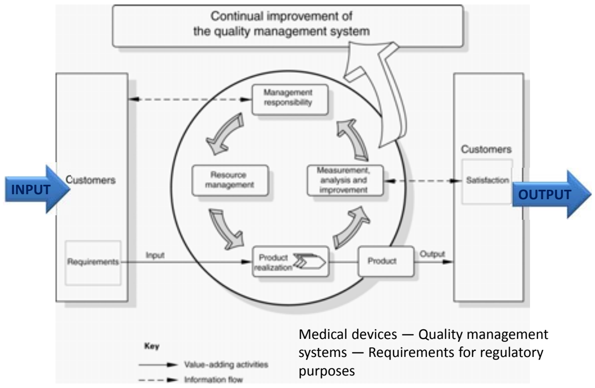
Peischl, B., Ferk, M. & Holzinger, A. 2015. The fine art of user-centered software development. Software Quality Journal, 23, (3), 509-536.

Remember: In medicine we have two different worlds …

Our central hypothesis: Information bridges this gap

Holzinger, A. & Simonic, K.-M. (eds.) 2011. ***Information Quality in e-Health.*** *Lecture Notes in Computer Science LNCS 7058, Heidelberg, Berlin, New York: Springer.*

Slide 12-10a ISO 13485:2003 Quality Management Process



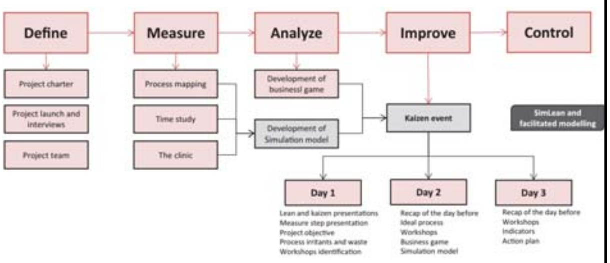
Medical devices — Quality management systems — Requirements for regulatory purposes

Slide 12-10b The origins: Kaizen

- Continuous improvement
- Making errors.
- Show errors!
- Learn from errors!!!
- Involve everybody
- Process oriented
- From small steps to big results

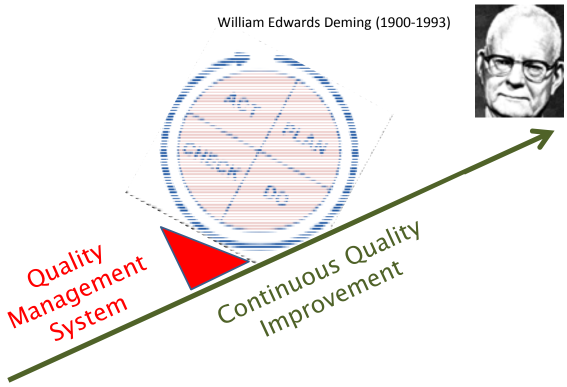
Masaaki, I. 1986. Kaizen: The Key to Japan's Competitive Success. New York: Random House.

Slide 12-10c The origins: Kaizen

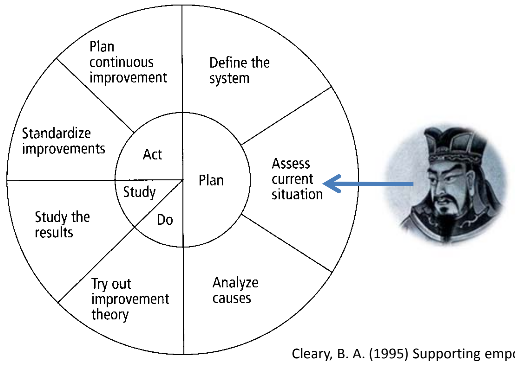


Baril, C., Gascon, V., Miller, J. & Cote, N. 2016. Use of a discrete-event simulation in a Kaizen event: A case study in healthcare. European Journal of Operational Research, 249, (1), 327-339.

Slide 12d Deming Wheel



Slide 12-11: Quality Improvement Cycle

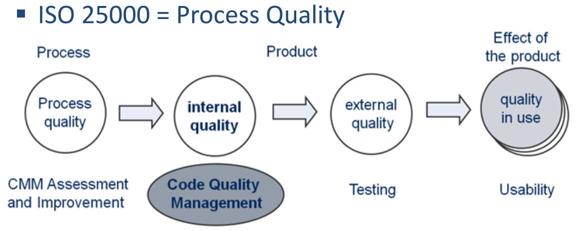


Cleary, B. A. (1995) Supporting empowerment with Deming's PDSA cycle. *Empowerment in Organizations, 3, 2, 34-39.*

Slide 12-12 Product vs. Process Quality

- ISO 9126 = Product Quality
- ISO 25000 = Process Quality

Capability Maturity Model (CMM)

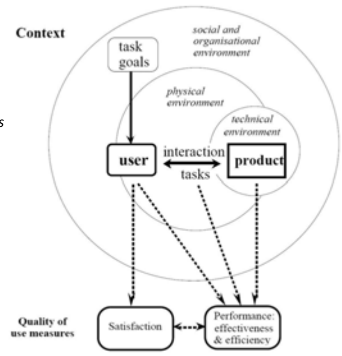


Plösch, R., Gruber, H., Hentschel, A., Körner, C., Pomberger, G., Schiffer, S., Saft, M. & Storck, S. (2008) The EMISQ method and its tool support-expert-based evaluation of internal software quality. *Innovations in Systems and Software Engineering, 4, 1, 3-15.*

Slide 12-13 The goal: Quality of Use = measured Usability

Holzinger, A., Stickel, C., Fassold, M. & Ebner, M. (2009) Seeing the System through the End Users' Eyes: Shadow Expert Technique for Evaluating the Consistency of a Learning Management System. In: *Lecture Notes in Computer Science (LNCS 5889). Heidelberg, Berlin, New York, Springer, 178-192.*

Bevan, N. (1995) Measuring Usability as Quality of Use. *Software Quality Journal, 4, **2, 115-130.***



Slide 12-14: ISO/IEC 9126-1 Software Product Quality

Functionality	Reliability	Efficiency
accuracy suitability interoperability security	maturity fault tolerance recoverability availability	time behaviour resource man. utilisation

Maintainability	Portability	Usability
analysability changeability stability testability	adaptability installability co-existence replaceability	understandability learnability operability attractiveness

Holzinger, A., Treitler, P. & Slany, W. 2012. Making Apps Useable on Multiple Different Mobile Platforms: On Interoperability for Business Application Development on Smartphones. In: Lecture Notes in Computer Science LNCS 7465. pp. 176-189.

Slide 12-15: Remember Medical workflows …

- *The quality of the work of physicians is heavily influenced by the usability of their available tools*

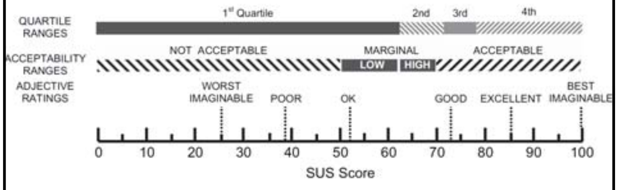
Holzinger, A., Geierhofer, R., Ackerl, S. & Searle, G. (2005). *CARDIAC@VIEW: The User Centered Development of a new Medical Image Viewer. Central European Multimedia and Virtual Reality Conference, Prague, Czech Technical University (CTU), 63-68.*

Slide 12-16: Comparison of Usability Engineering Methods

	Inspection Methods			Test Methods		
	Heuristic Evaluation	Cognitive Walkthrough	Action Analysis	Thinking Aloud	Field Observation	Questionnaires
Applicably in Phase	all	all	design	design	final testing	all
Required Time	low	medium	high	high	medium	low
Needed Users	none	none	none	3+	20+	30+
Required Evaluators	3+	3+	1-2	1	1+	1
Required Equipment	low	low	low	high	medium	low
Required Expertise	medium	high	high	medium	high	low
Intrusive	no	no	no	yes	yes	no

Holzinger, A. (2005) Usability engineering methods for software developers. *Communications of the ACM, 48, 1, 71-74.*

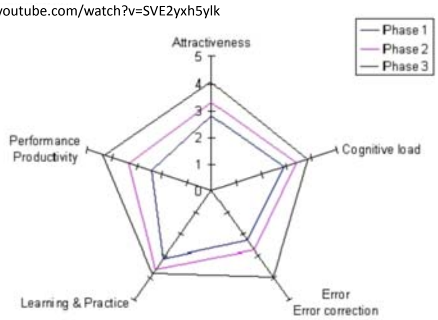
Slide 12-17: The System Usability Scale (SUS)



Bangor, A., Kortum, P. T. & Miller, J. T. (2008) An empirical evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction, 24, 6, 574-594.*

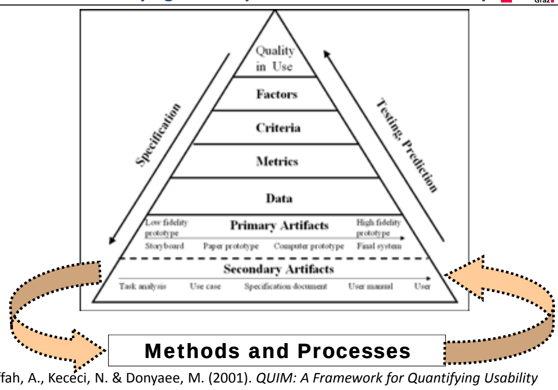
Slide 12-18: Software Usability Measurement Inventory SUMI

A funny video about SUMI can be found here:
http://www.youtube.com/watch?v=SVE2yxh5ylk



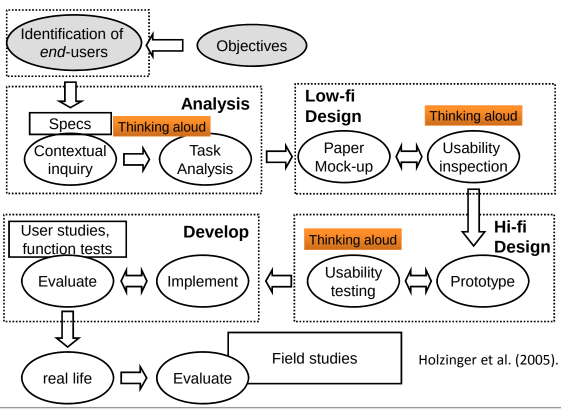
Kosec, P., Debevc, M. & Holzinger, A. 2009. Towards Equal Opportunities in Computer Engineering Education: Design, Development and Evaluation of Video-based e-Lectures. International Journal of Engineering Education (IJEE), 25, (4), 763-771.

Slide 12-19 Quantifying Usability Metrics in Software Quality



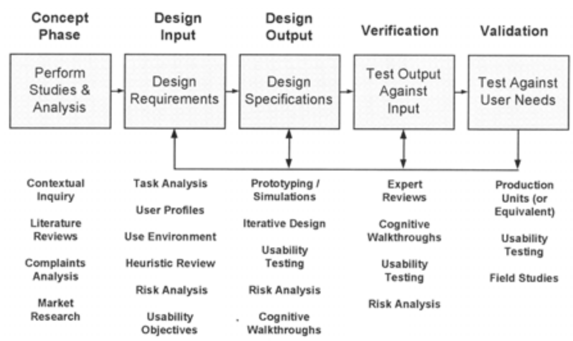
Seffah, A., Kececi, N. & Donyaee, M. (2001). *QUIM: A Framework for Quantifying Usability Metrics in Software Quality Models. APAQS'01, Hong Kong, 311-318.*

Slide 12-20 User Centred Design and Development (UCD)



Holzinger et al. (2005).

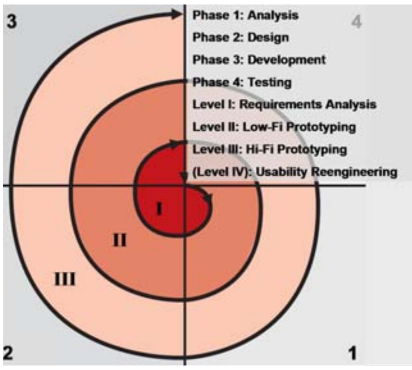
Slide 12-21: Remember the big picture: UCD Process



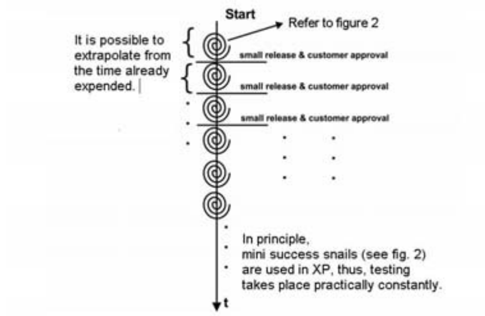
Wiklund, M. E. & Wilcox, S. B. (2005) *Designing Usability into Medical Products. Boca Raton et al., Taylor & Francis.*

Slide 12-22 The power of iteration: A UCD spiral

Holzinger, A. (2004) Application of Rapid Prototyping to the User Interface Development for a Virtual Medical Campus. *IEEE Software, 21, 1, 92-99.*

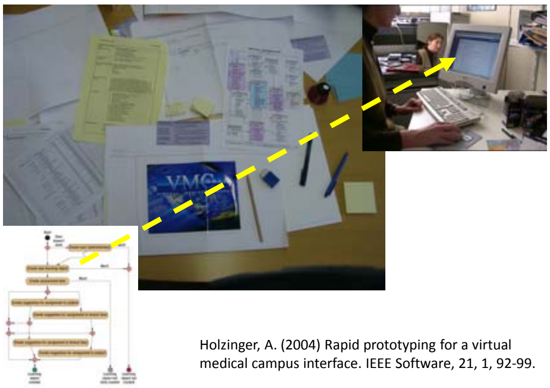


Slide 12-23: Agility: Make the UCD spirals as small as possible



Holzinger, A. & Slany, W. (2006) XP + UE -> XU Praktische Erfahrungen mit eXtreme Usability. *Informatik Spektrum, 29, 2, 91-97.*

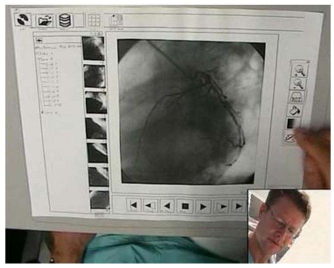
Slide 12-24 Rapid Prototyping – Paper Mock-ups



Holzinger, A. (2004) Rapid prototyping for a virtual medical campus interface. IEEE Software, 21, 1, 92-99.

Slide 12-25 Insight into the end user: Thinking aloud

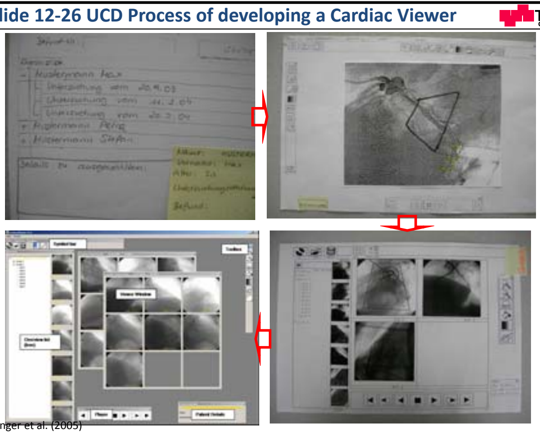
- Important to implement this method as early as possible in the software development process
- the later that understanding of the user's behaviour is gained, the more improbable it is that these can still be integrated into the development.



Brown, S. & Holzinger, A. (2008) Low cost prototyping: Part 1, or how to produce better ideas faster by getting user reactions early and often. In: Abuelmaatti, O. & England, D. (Eds.) *Proceedings of HCI 2008. Liverpool: John Moores University (UK), British Computer Society, 213–214.*

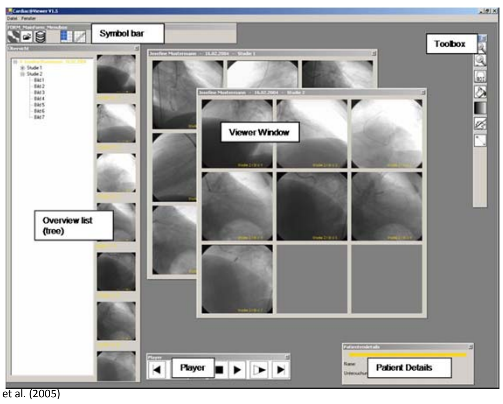
Holzinger, A. & Brown, S. (2008) Low cost prototyping: Part 2, or how to apply the thinking-aloud method efficiently. In: Abuelmaatti, O. & England, D. (Eds.) *Proceedings of HCI 2008. Liverpool: John Moores University (UK), British Computer Society, 217–218.*

Slide 12-26 UCD Process of developing a Cardiac Viewer



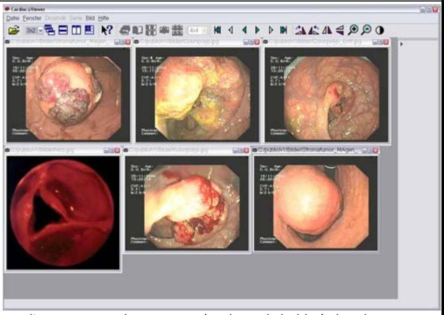
Holzinger et al. (2005)

Slide 12-27 Hi-Fi Prototype allows low-level interaction



Holzinger et al. (2005)

Slide 2-28 Validation & Verification to check quality



Validation = is a (external) quality process to demonstrate (to the stakeholder) *that* the system complies with the original specifications;
Verification = is a (internal) quality process, used to evaluate whether and *to what extent* the system complies with the original specifications;

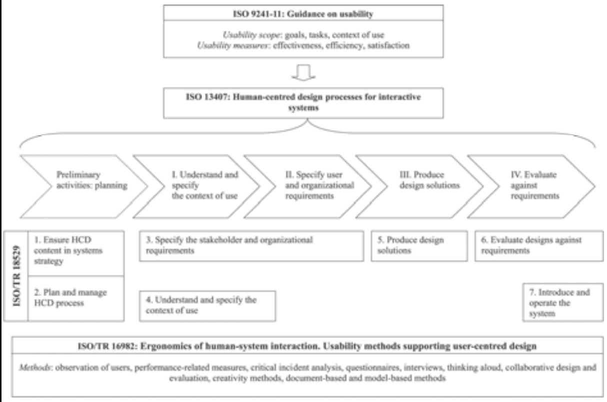
Holzinger et al. (2005)

Slide 12-29 ISO 13407 Human-Centred Design (1/2)

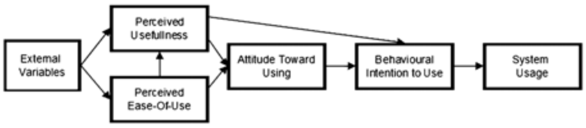
Title	ISO 13407 *Human-centred design processes for interactive systems*
Date	July 1999
Scope	Guidance on human-centred design activities throughout the lifecycle of interactive computer-based systems.
Contents	The rationale for a user-centred design process. A description of the four core principles of human-centred design. Planning of the user-centred design process. Description of the four activities which should take place during a system development process. A listing of current process and product standards for user-centred design.
Purpose	ISO 13407 aims to help those responsible for managing hardware and software design processes to identify and plan effective and timely user-centred design activities. It complements existing design approaches and methods.
Audience	Those managing the design process. All parties involved in human-centred system development, including the end-users of systems, are expected to find the standard relevant.
Requirements	Any development process which claims to have met the recommendations in ISO 13407 shall specify the procedures used, information collected and use made of the results.

Earthy, J., Jones, B. S. & Bevan, N. (2001) The improvement of human-centred processes - facing the challenge and reaping the benefit of ISO 13407. *International Journal of Human-Computer Studies, 55, 4, 553-585.*

Slide 12-30 ISO 13407 Human-Centered Design (2/2)



Slide 12-31 Technology Acceptamce Model 75* > 89** > 11



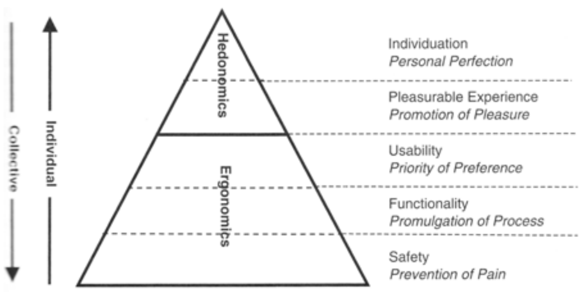
It was experimentally proved that the acceptance is related to a factor, which is called **PET (Previous Exposure to Technology)**

Holzinger, A., Searle, G. & Wernbacher, M. 2011. The effect of Previous Exposure to Technology (PET) on Acceptance and its importance in Usability Engineering. Springer Universal Access in the Information Society International Journal, 10, (3), 245-260.

*) Davis, F. D. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 13, (3), 319-339.

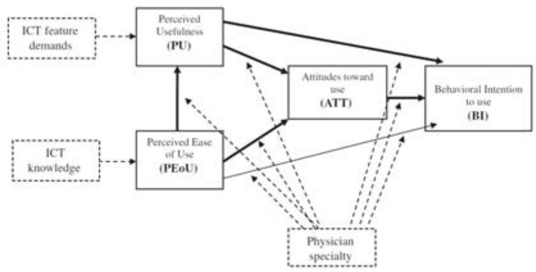
**) Fishbein, M. & Ajzen, I. 1975. Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research, Reading (MA), Addison-Wesley.

Slide 12-32 Ergonomics versus Hedonomics



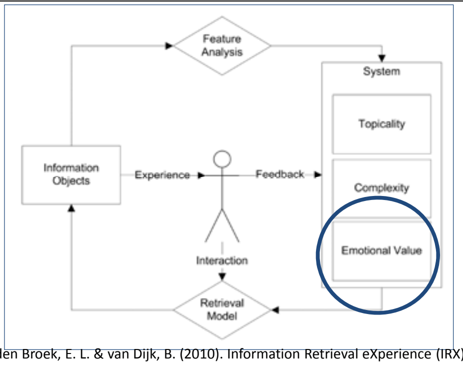
Helander, M. G. & Khalid, H. M. (2006) Affective and Pleasurable Design. In: Salvendy, G. (Ed.) *Handbook of Human Factors and Ergonomics, Third Edition. Hoboken (NJ), Wiley.*

Slide 12-33 Technology Acceptance in the clinical context



Melas, C. D., Zampetakis, L. A., Dimopoulou, A. & Moustakis, V. (2011) Modeling the acceptance of clinical information systems among hospital medical staff: An extended TAM model. *Journal of Biomedical Informatics, 44, 4, 553-564.*

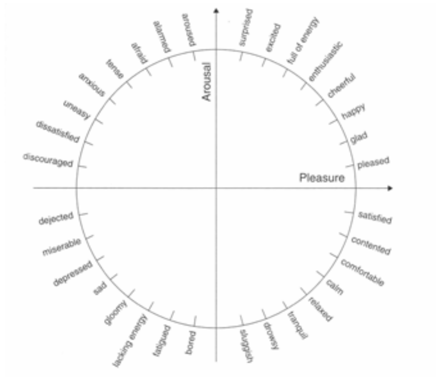
Slide 12-34 Example: Information Retrieval Experience



Sluis, F., van den Broek, E. L. & van Dijk, B. (2010). Information Retrieval eXperience (IRX): Towards a Human-Centered Personalized Model of Relevance. Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on, 322-325.

Slide 12-35 Example: Emotion 2-D measurement scale

Helander, M. G. & Khalid, H. M. (2006) Affective and Pleasurable Design. In: Salvendy, G. (Ed.) *Handbook of Human Factors and Ergonomics, Third Edition. Hoboken (NJ), Wiley.*



Slide 12-36 How to measure emotions?

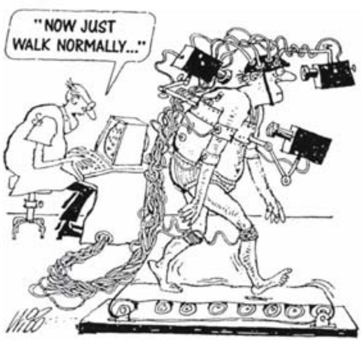
- **Neuro-physiological,** e.g. brain activity, pulse rate, blood pressure, skin conductance, etc.
  - Can detect short-term changes not measurable by other means; Reliance on non-transparent, invasive sensors; can reduce people's mobility, causing distraction of emotional reactions; prone to noise due to unanticipated changes in physiological characteristics; inability to map data to specific emotions; require expertise and the use of special, often expensive, equipment
- **Observation,** e.g. facial expressions; speech; gestures Use of unobtrusive techniques for measuring emotion; cross-cultural universals
  - Can not perform context dependent interpretation of sensory data; highly dependent on environmental conditions (illumination, noise, etc.); some responses can be faked; recognizes the presence of emotional expressions, not necessarily emotions
- **Self-reporting,** e.g. questionnaire, diary; interview;
  - High correlation to neurophysiological evidence; unobtrusive; straightforward and simple – do not require the use of special equipment; Rely on the assumption that people are aware of and willing to report their emotions; subject to the respondent's bias; results of different studies might not be directly comparable

Lopatovska, I. & Arapakis, I. (2011) Theories, methods and current research on emotions in library and information science, information retrieval and human–computer interaction. Information Processing & Management, 47, 4, 575-592.

Slide 12-37 Example methods for measuring emotion

- Subjective measures -> Kansei Engineering, Semantic scales (e.g. Nagamachi (2001), Helander & Tay (2003)); Experience sampling method (e.g. Larson & Csikszentmihayi (1983); Affect Grid (e.g. Russel et al. (1989), Warr (1999); MACL Checklist (e.g. Nowlis & Green (1957)); PANAS Scale (e.g. Watson et al. (1988)); Philips questionnaire (e.g. Jordan (2000));
- Objective Measures -> Facial action coding system (e.g. Ekman (1982); Maximally discriminative affect coding system (e.g. Izard (1979); Facial electromyography (e.g. Davis et al. (1995);
- Psychogalvanic measures -> Galvanic skin response (e.g. Larson & Fredrickson (1999), Wearable sensors (e.g. Picard (2000);
- Performance measures -> Judgment task involving probability estimates (e.g. Katelaar (1989); Lexical decision task (e.g. Challis & Krane (1988), Niedenthal & Setterlund (1994)

Slide 12-38 Problem: Obtrusiveness of measuring



Ouwerkerk, M., Pasveer, F. & Langereis, G. (2008) Unobtrusive Sensing of Psychophysiological Parameters: Some Examples of Non-Invasive Sensing Technologies. In: Westerink, J. H. D. M. (Ed.) *Probing Experience. Heidelberg, Berlin, New York, Springer, 163-193.*

Evaluation

Remember: Traditional Programming vs Machine Learning

Traditional Programming

Data → Computer → Output
Program → Computer

Machine Learning = Learning from Data

Data → Computer → Program
Output → Computer

Slide 12-39 Occam's Razor: take the simplest alternative

Occam's Razor: No more things should be presumed to exist than are absolutely necessary, i.e., the fewer assumptions an explanation of a phenomenon depends on, the better the explanation.

(William of Occam)

Nunquam ponenda est pluralitas sin necesitate," which, approximately translated, means Entities should not be multiplied beyond necessity

Domingos, P. 1999. The role of Occam's razor in knowledge discovery. Data mining and knowledge discovery, 3, (4), 409-425.

Slide 12-40 NFL-Theorem

Wolpert, D. H. & Macready, W. G. 1997. No free lunch theorems for optimization. IEEE Transactions on Evolutionary Computation, 1, (1), 67-82.

Slide 12-41 Performance Measures (selection)

- Scalability
- Predictive accuracy = Hit rate
- Weighted (cost-sensitive) accuracy
- Speed (on model building and predicting)
- Robustness (one weakness in iML-approach)
- Precision/Recall (F-Measure, Break Even Point)
- Area under the ROC (see next slides)

Japkowicz, N. & Shah, M. 2011. Evaluating learning algorithms: a classification perspective, Cambridge University Press.

FYI: Datasets for benchmarking purposes

- There are many datasets for testing machine learning algorithms, just some examples:
- https://www.kaggle.com
- http://archive.ics.uci.edu/ml/datasets.html (UCI Machine Learning Repository)
- http://image-net.org
- http://yann.lecun.com/exdb/mnist (handwritten digit database)
- https://data.medicare.gov/

http://hci-kdd.org/open-data-sets/

Accuracy

- **Question: is 99% accuracy good?**
- **Answer: It depends on the problem!**

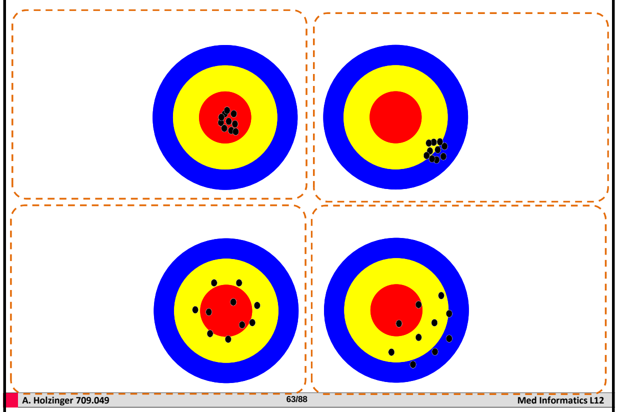
Please always remember these four terms:

- **Accuracy** = error rate of correct/incorrect predictions made by the model over a data set (cf. coverage).
- **Precision** = precision (positive predictive value) is the fraction of retrieved instances that are relevant, while **Recall** (aka sensitivity) is the fraction of relevant instances that are retrieved
- **Reliability** = basically the "consistency" or "repeatability"
- **Validity** = generally, to get valid conclusions

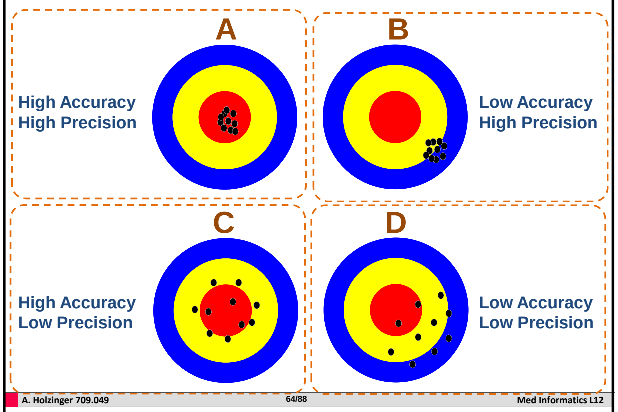
Accuracy **Validity**

Precision **Reliability**

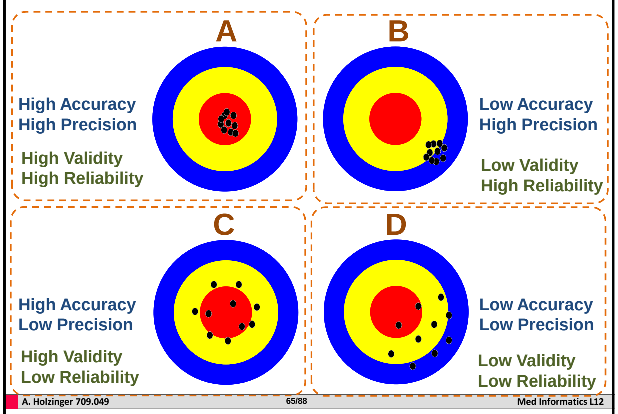
Accuracy vs Prediction: Examples



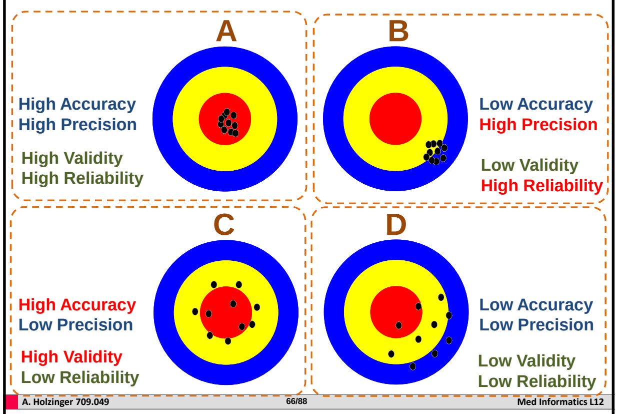
Accuracy vs. Precision



Accuracy vs. Precision



Accuracy vs. Precision



Please remember:

		True Class: Positive	True Class: Negative
Predicted Class	Positive	True Positive Count (TP)	False Positive Count (FP)
	Negative	False Negative Count (FN)	True Negative Count (TN)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

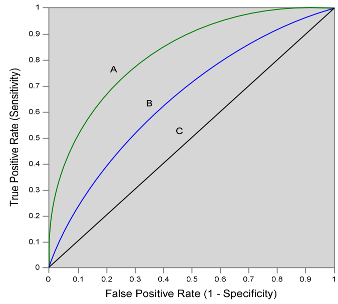
$$True\ Positive\ Rate = \frac{TP}{TP + FN}$$

$$True\ Negative\ Rate = \frac{TN}{TN + FP}$$

$$Precision = \frac{TP}{TP + FP} \qquad Recall = \frac{TP}{TP + FN}$$

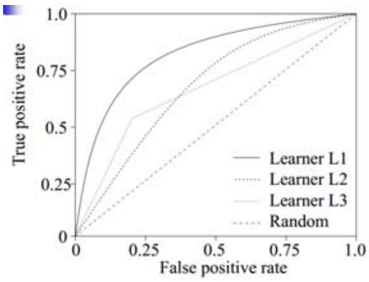
Turban, E., Sharda, R., Delen, D. & Efraim, T. 2007. Decision support and business intelligence systems, Pearson Education.

Again the ROC Curve



Bradley, A. P. 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition,* 30, (7), 1145-1159.

Examples



For a detailed explanation refer to: Fawcett, T. 2006. An introduction to ROC analysis. Pattern recognition letters, 27, (8), 861-874.

Future Outlook

- Classification and Prediction
- Decision Tree
- Support Vector Machine (SVM)
- Evaluation (Accuracy of Classification Model)

A last word …

Hans Holbein d.J., 1533, The Ambassadors, London: National Gallery

Lopez-Paz, D., Muandet, K., Schölkopf, B. & Tolstikhin, I. 2015. Towards a learning theory of cause-effect inference. Proceedings of the 32nd International Conference on Machine Learning, JMLR, Lille, France.

https://www.youtube.com/watch?v=9KiVNIUMmCc

Thank you!

Sample Questions (1/2)

- What does Total Workplace Usability include and why is this important to enhance quality?
- What are the key measurable concepts of usability?
- Please describe the overall UCD Process from concept to validation!
- Which are the corresponding quality factors of safety critical medical systems?
- What does the EU directive 93/42 Medical Device Directive (MDD) describe?
- Why is now for system developers/providers usability not only relevant but also mandatory?
- What does ISO 14971:2007 describe?
- Please describe the principles of the quality improvement cycle!
- What does ISO 13407 describe?
- Please describe the three most important Usability Inspection Methods!

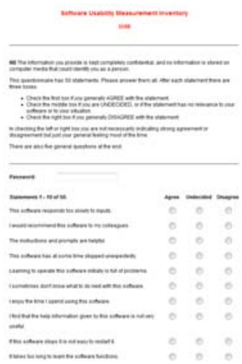
Sample Questions (2/2)

- Please describe the three most important Usability Test Methods!
- How would you apply the System Usability Scale (SUS)?
- What is the difference between Lo-Fi and Hi-Fi Prototyping?
- What is the advantage of a paper mock-up?
- How to you perform a Thinking aloud test?
- What is the difference between Hedonomics and Ergonomics?
- Why is emotion an important aspect to consider?
- Which possibilities do you have to measure emotion?
- What is the disadvantage of Neuro-physiological methods?
- What is the difference between Validation and Verification?
- Why do we speak of an end-user? Why is just "user" not sufficient?
- What is the purpose of a quality audit?

Some useful links (1)

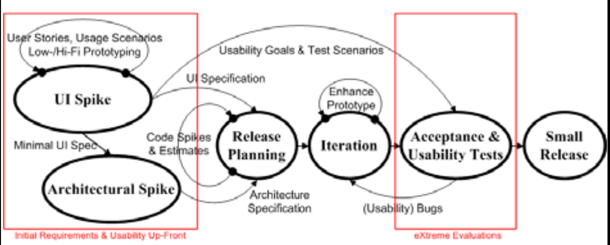
- http://www.measuringusability.com/sus.php (Measuring Usability with the System Usability Scale (SUS))
- http://sumi.ucc.ie (Software Usability Measurement Inventory (SUMI))
- http://www.gesetze-im-internet.de/mpg/index.html (Gesetz über Medizinprodukte - Deutschland)
- http://www.jusline.at/Medizinproduktegesetz_%28MPG%29.html (Medizin Produkte Gesetz, MPG – Österreich)
- http://www.iso.org/iso/iso_9000_selection_and_use.htm (Selection and use of the ISO 9000 family of standards)
- https://www.dsk.gv.at/site/6274/default.aspx (Österreichische Datenschutzkommission, Austrian Data Protection Commission)
- http://www.ethikkommissionen.at (Ethical Commissions in Austria)
- http://iaidq.org (The International Association for Information and Data Quality (IAIDQ))
- http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31993L0042:EN:HTML (Council Directive 93/42/EEC of 14 June 1993 concerning medical devices)
- http://ec.europa.eu/health/medical-devices/index_en.htm (European Commission, Public Health, Medical Device Act)
- http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_tc_browse.htm?commid=54960 (ISO Standards Technical Committee TC 215 Health Informatics)
- http://www.iso.org/iso/hot_topics.htm (Hot Topics Section of the International Standardization Organisation)
- http://www.iso.org/iso/pressrelease.htm?refid=Ref1304 (Protecting integrity and privacy of electronic medical records with new ISO guidelines)

Appendix: Software Usability Measurement Inventory



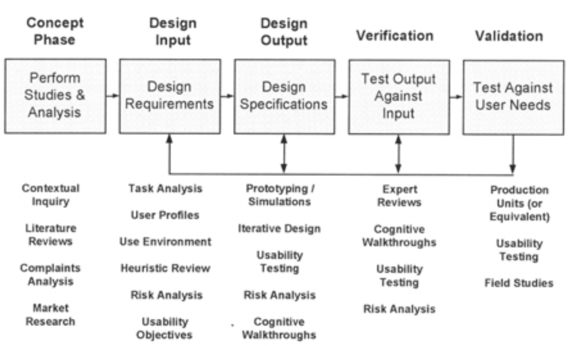
http://sumi.ucc.ie/en/

Appendix: Agile Process Model



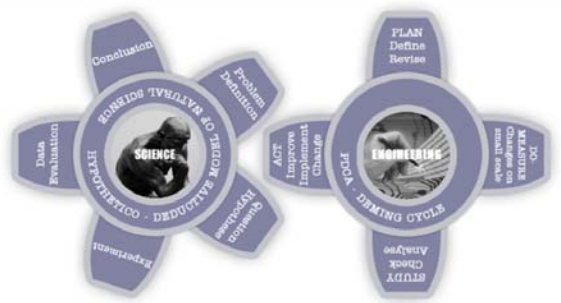
Memmel, T., Reiterer, H. & Holzinger, A. (2007) Agile Methods and Visual Specification in Software Development: a chance to ensure Universal Access. *Coping with Diversity in Universal Access, Research and Development Methods in Universal Access, Lecture Notes in Computer Science (LNCS 4554). Berlin, Heidelberg, New York, Springer, 453-462.*

Slide 12-3: The big picture: UCD Process



Wiklund, M. E. & Wilcox, S. B. (2005) *Designing Usability into Medical Products. Boca Raton et al., Taylor & Francis.*

HCI - Combine Science and Engineering



http://www.hci4all.at

Comparison of Usability Engineering Methods

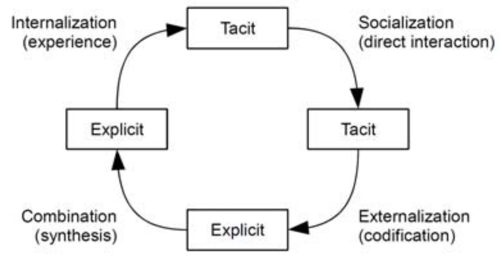
	Inspection Methods			Test Methods		
	Heuristic Evaluation	Cognitive Walkthrough	Action Analysis	Thinking Aloud	Field Observation	Questionnaires

Holzinger, A. (2005) Usability engineering methods for software developers. *Communications of the ACM, 48, 1, 71-74.*

Criteria	Efficiency	Effectiveness	Satisfaction	Productivity	Learnability	Safety	Trustfulness	Accessibility	Universality	Usefulness
Time behavior	+			+						
Resource utilization	+			+						+
Attractiveness			+						+	
Likeability			+							
Flexibility		+	+					+	+	+
Minimal action	+		+		+			+		
Minimal memory load	+		+		+			+	+	+
Operability	+		+				+	+		+
User guidance			+		+			+	+	
Consistency		+			+	+		+	+	
Self-descriptiveness					+		+	+	+	
Feedback	+	+							+	+
Accuracy		+				+				+
Completeness		+				+				
Fault-tolerance						+	+			+
Resource safety						+				
Readability								+	+	
Controllability							+	+	+	+
Navigability	+	+					+	+	+	
Simplicity					+			+	+	
Privacy							+		+	+
Security						+	+			+
Insurance						+	+			
Familiarity					+		+			
Loading time	+			+					+	+

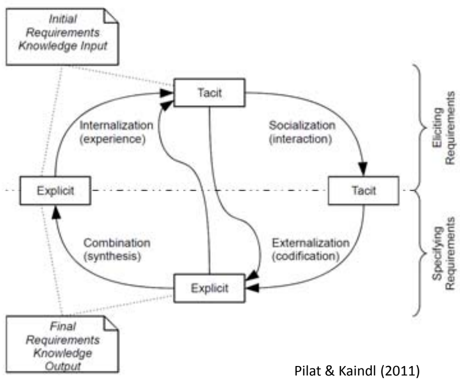
(Factors: Efficiency, Effectiveness, Satisfaction, Productivity, Learnability, Safety, Trustfulness, Accessibility, Universality, Usefulness)

Remember: Cyclic View of Nonaka's Spiral of Knowledge



Pilat, L. & Kaindl, H. (2011) A knowledge management perspective of requirements engineering. *Fifth International Conference on Research Challenges in Information Science (RCIS). 1-12.*

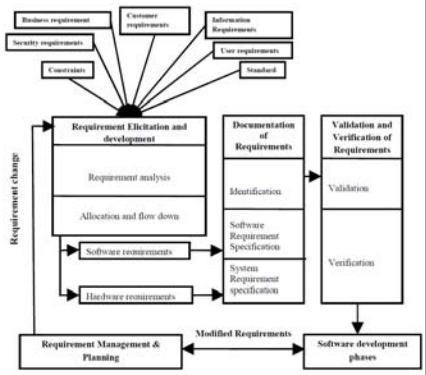
Spiral of Requirements Knowledge



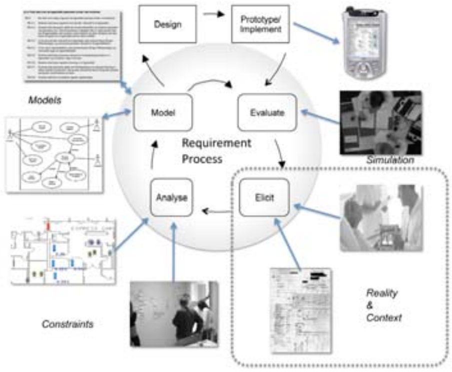
Pilat & Kaindl (2011)

Example: Requirement Engineering Process Model

Pandey, D., Suman, U. & Ramani, A. K. (2010) An Effective Requirement Engineering Process Model for Software Development and Requirements Management. *International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom). 287-291.*



Elicitation in the requirements process in the health domain



Nytro, O., Sorby, I. D. & Karpati, P. (2009) Query-based requirements engineering for health care information systems: Examples and prospects. *ICSE Workshop on Software Engineering in Health Care. 62-72.*

Actors and Information Categories

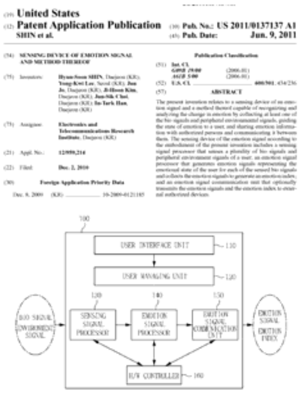
Category	*Value*
Human	Patient Next of kin Ward secretary Physician Nurse
Paper based	Patient chart Patient record Ward list (patient summaries) Patient information (discharge) Schemas ICD-10 code overview Prescription Physicians' Desk Reference (PDR) Appointment scheduling book Personal notes
Electronic	Electronic patient record (EPR) Patient administrative system (PAS) Physicians' Desk Reference (PDR) Personal digital assistant (PDA) PACS/RIS (Picture archive & comm.

Category	*Value*
Patient information	Biographical data (BIO) Family/social history (FAMSOS) Resum/overview of patient
Past	Allergies Reason for referral (REASON) Previous illnesses (PREVILL)
Present	Diagnosis (D) Assessment Blood tests/results (BLOOD) Electrocardiogram (ECG) Examination Progress and treatment (PROGTREAT) Findings and examination results (FINDEX) Medication administration (MED)
Future	Procedure Plan for investigation (PFI) Plan for treatment (PFT) Medications (prescriptions) (MED) Info. to patient/next of kin Prescription Requisition Discharge Follow-up

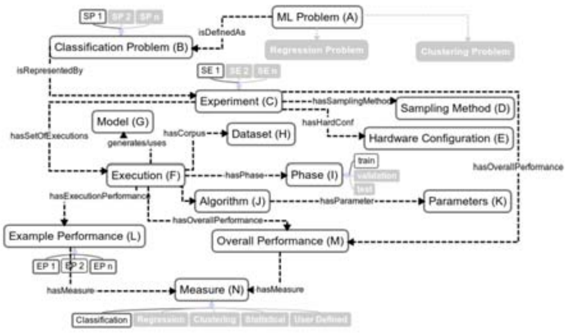
Nytro, Sorby & Karpati (2009)

Example Patent Application A1

US Kind Codes: Before January 2001 patents had the label A and patent applications the label B1, B2, …; however, since January 2001, US Patents are labelled differently: A1 is the first patent application, A2 the second, etc., whereas B1, B2, … are the granted patens! X-documents are problematic, because every Xdocument is detrimental for any further patent application in the area of the X-document!



Holzinger, A. (2010) *Process Guide for Students for Interdisciplinary Work in Computer Science/Informatics. Second Edition. Norderstedt, BoD.*



Esteves, D., Moussallem, D., Neto, C. B., Soru, T., Usbeck, R., Ackermann, M. & Lehmann, J. MEX vocabulary: a lightweight interchange format for machine learning experiments. Proceedings of the 11th International Conference on Semantic Systems, 2015. ACM, 169-176.