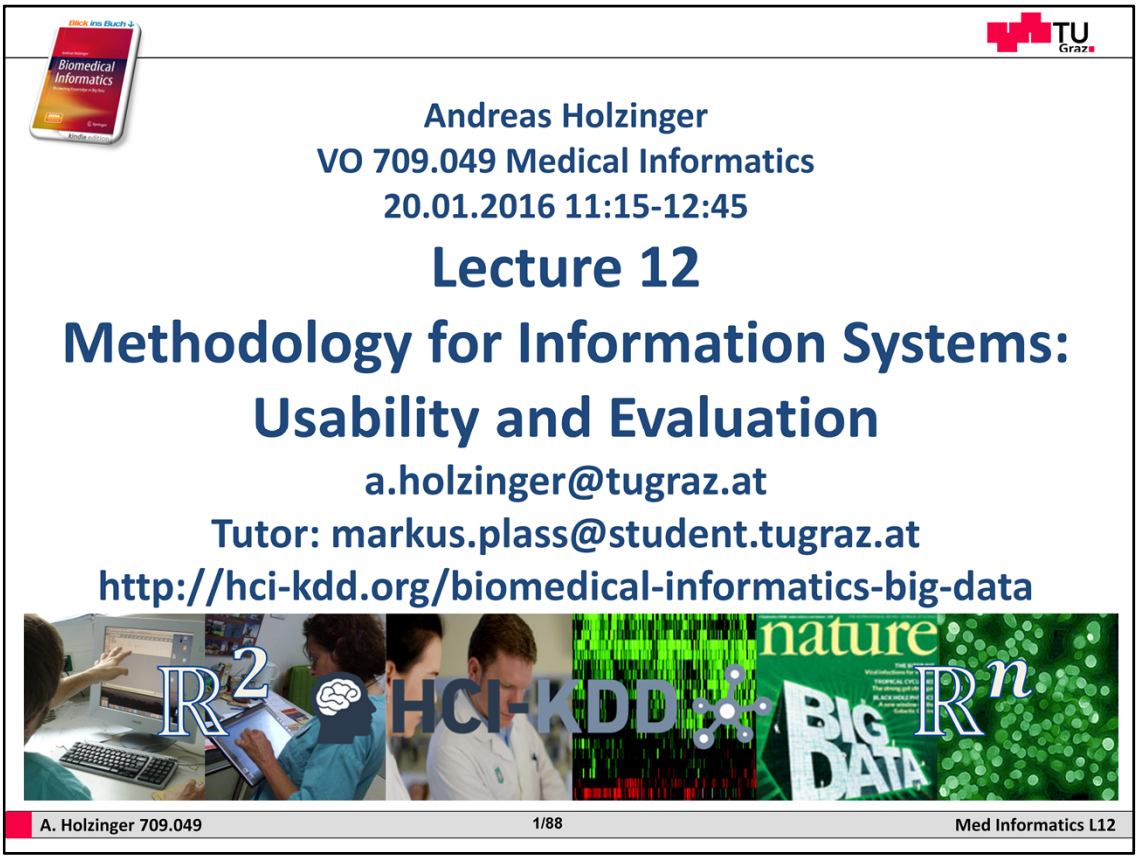


Status as of 19.01.2016 09:00

Dear Students, welcome to the 12th and last lecture of our course, where we will deal with some methods of usability and evaluation.

Please remember from the last lecture that in **all** biomedical applications, privacy, data protection, safety and security issues are mandatory!

Please always be aware of the definition of biomedical informatics (Medizinische Informatik):
Biomedical Informatics is the inter-disciplinary field that studies and pursues the effective use of biomedical data, information, and knowledge for scientific inquiry, problem solving, and decision making, motivated by efforts to improve human health and well-being.

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 ☺

A. Holzinger 709.049 3/88 Med Informatics L12

Evaluation (Bewertung) = can be formative or summative
Benchmarking (Leistungsvergleich) = typically measuring according to: cost, time and quality.

At the end of this last lecture you …
… understand the concepts of usability and the importance of usability engineering for medical information systems;
… are aware that medical software is now included within the Medical Device Act (Medizin Produkte 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;

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

Quality is the central topic of this lecture, particularly information quality, as it applies to all issues in any biomedical informatics application.

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.

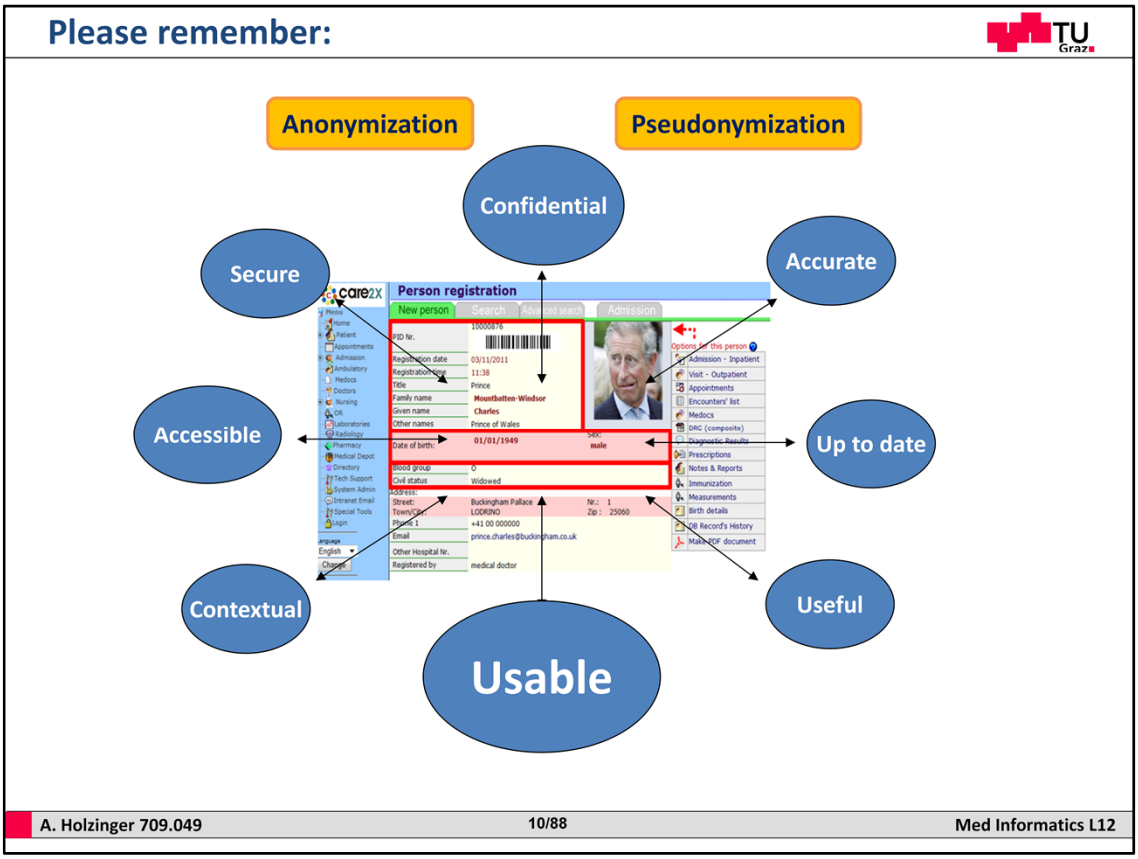
[2] Thimbleby, H. 2007. User-Centered Methods Are Insufficient for Safety Critical Systems. Lecture Notes in Computer Science (LNCS 4799). Springer, pp. 1-20.

[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.

- Usability is still underestimated in health applications
- User-Centred Designs are rarely applied in medical information systems
- Evaluation is still a small part in medical information systems research

Usability is still underestimated in the design and development of applications for medicine and health care although they are proven to be often a matter of life or death. Jakob Nielsen reported this very impressive in his blog, under the title: "How to kill patients through bad design", http://www.nngroup.com/articles/medical-usability
Medical systems have provided many well-documented killer designs, such as the radiation machines that fried six patients because of complex and misleading operator consoles. What's less known is that usability problems in the medical sector's good old-fashioned office automation systems can harm patients just as seriously as machines used for treatment. A further problem is that traditional approaches of HCI are essential, but they are unable to cope with the complexity of typical modern interactive devices in the safety critical context of medical devices. The broad scale of typical devices means that conventional user-centered approaches, while still necessary, are insufficient to contribute reliably to safety related interaction issues (Thimbleby, 2007).

Algorithm performance evaluation is so entrenched in the machine learning community that one could call it an addiction.



Requirements of an electronic patient record
Remember the requirements to a patient record from the viewpoint of ensuring privacy: The patient data must be confidential, secure and safe, while at the same time must be usable, useful, accurate, up-to-date and accessible.

Security issues are crucial in a number of machine learning applications, especially in scenarios dealing with human activity rather than natural phenomena (e.g., information ranking, spam
detection, malware detection, etc.). In such cases, learning algorithms may have to cope with manipulated data aimed at hampering decision making. Although some previous work addressed the
issue of handling malicious data in the context of supervised learning, very little is known about
the behavior of anomaly detection methods in such scenarios.

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.*

A. Holzinger 709.049 11/88 Med Informatics L12

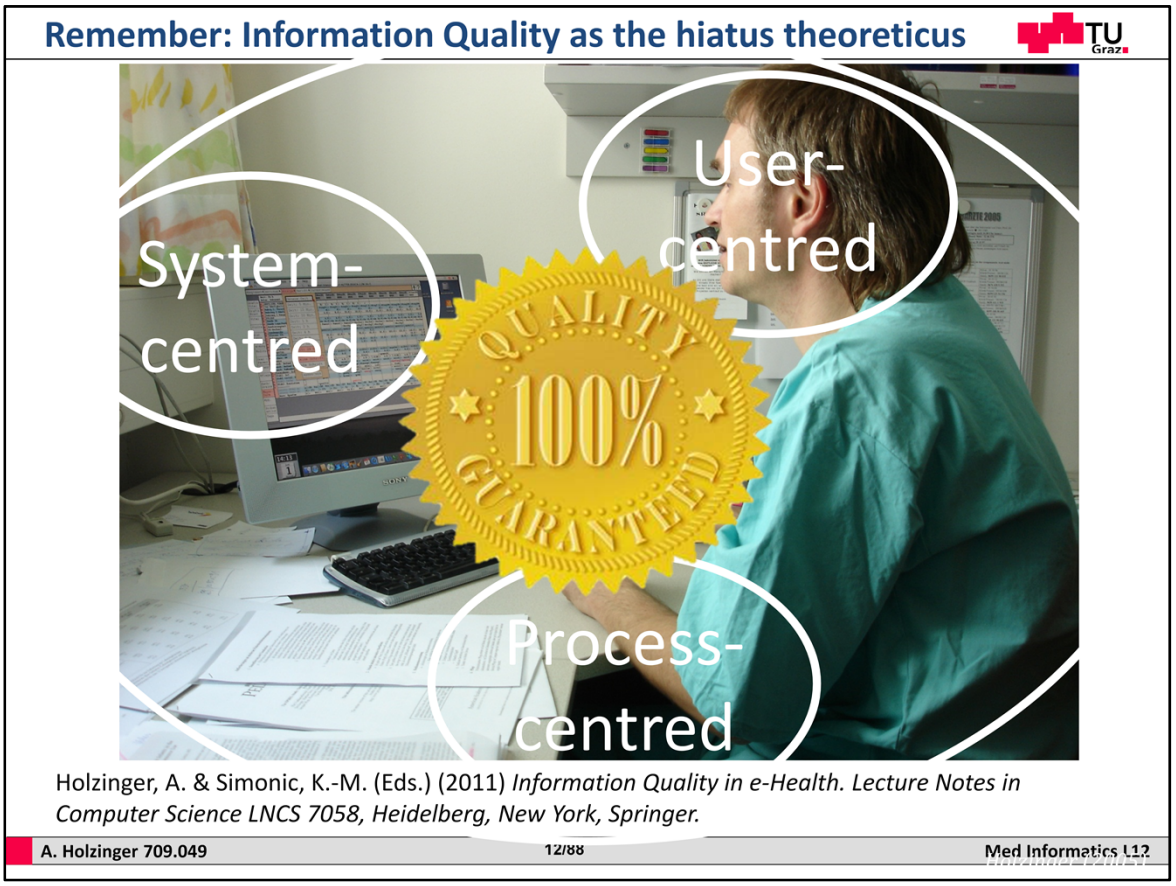
In this slide we see a typical usability survey setting: Experiments with end-users in real-world settings in the Hospital often show that the whole workflow, along with human factors of the workplace, including hardware usability, e.g. seating and environmental issues must be considered. To ensure a total acceptance of a system, satisfaction and most of all the reduction in **time to perform task** is essential. To reach that, all issues of the whole workflow must be considered which is called Total Workplace Usability (TWU) (Holzinger & Leitner, 2005). The issues to consider include:

• On-site training to familiarize the end-users with the tools having one big goal in mind: reduce time to perform a task. Time is precious for the medical doctors and each time saved is also a cost saved for the hospital.

• Electronic Tutoring to assist part-time end-users to find their way through the workflow quickly and to help them to solve their problems rapidly. Benefit: once developed, it would constantly assist many end-users at all workplaces, independent of end-users training efforts;

• Customizing and proper adaptation of third-party Software; in practice most tools separately used to the clinical workplace, especially when rarely used, cause a serious effort by the end-users.

• Ergonomic Aspects of the workplace include: proper distance to the screen, correct table height, proper chair (easily adjustable for height, to enable the user to quickly achieve a comfortable angle from eyes to the screen), proper mouse pad location and working space. Benefit: End-users can concentrate and feel more comfortable during the strenuous information processing process; results are immediate in lower task performance time and consequently a better output.

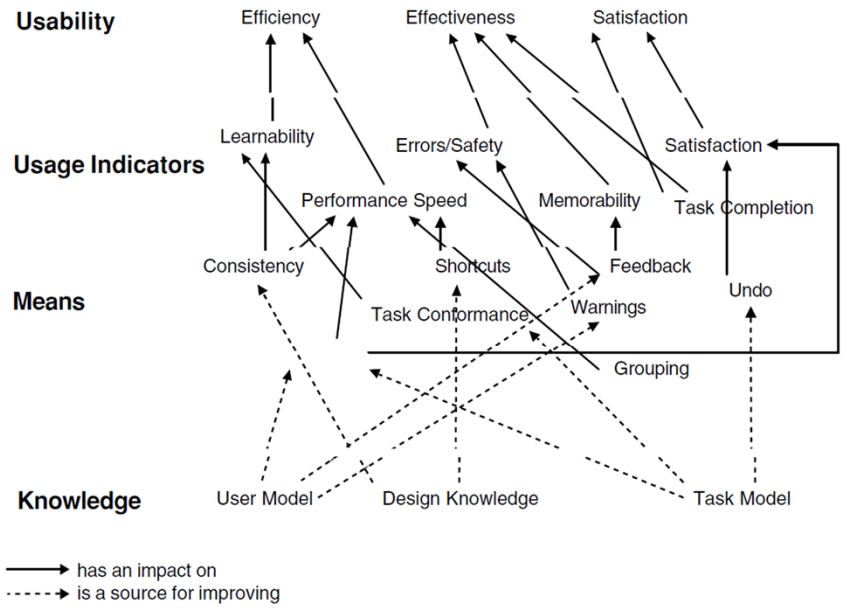


Medical Information Systems of today are highly sophisticated; however, while we have seen that computer performance has increased exponentially, the human cognitive evolution cannot advance at the same speed. Consequently, the focus on interaction between human and computer is of increasing importance. The daily actions of medical professionals within the context of their clinical work must be the central concern of any innovation. Just surrounding and supporting them with new and emerging technologies is not sufficient if these increase rather than decrease the workload.

Quality, actually, is a term which both Medicine as well as Informatics accept as an important issue (Holzinger & Simonic, 2011), and must include the user-centred (human), the system-centred (computer) and process-centered (interaction) view.

Total Quality Management (TQM) provides a useful, simple yet important definition of quality: "consistently meeting customer's expectations." (Fisher, Lauría & Chengalur-Smith, 2012). However, in medicine, this goal is not easy to accomplish, due to a number of problems, see Slide 1-45.

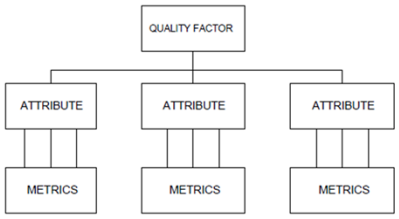
Slide 12-3: A framework for understanding usability



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.*

Usability can be defined via a combination of efficiency, effectiveness and satisfaction: Each of so-called **usage indicators** contributes to the aspects of the higher level, e.g. low error rate increases effectiveness; good performance indicates good efficiency, etc. The indicators are measured using a set of metrics. One level lower is the level of means, which can be used in "heuristics" for improving the usage indicators and are not goals by themselves, e.g. consistency may have a positive effect on learnability, as warnings may reduce errors. On the other hand, high adaptability may have a negative effect on memorability while having a positive effect on performance. In order to find optimal levels for all means, the designer has to apply the 3 knowledge domains: humans, design, and task. For example, design knowledge such as guidelines should include how changes in use of the means affect the usage indicators (Veer & Welie, 2004).

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.*

A. Holzinger 709.049 14/88 Med Informatics L12

Quality can be seen as the key success factor and if we look at our classic system characteristics we can determine six characteristics with corresponding quality factors, as seen in this slide. From top to bottom it includes: reliability, correctness and verifiability; security, efficiency, portability, usability and expandability – and the nature of the system dictates the prioritizing of the features (Cosgriff, 1994)

Slide 12-5: ISO Standards for Healthcare

ISO/TC 215 Health informatics

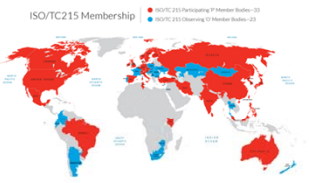
About | Contact details | Structure | Liaisons | Meetings | Tools

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

A. Holzinger 709.049 15/88 Med Informatics L12

In Lecture 3 we have already heard about the advantages and disadvantages of standardization. Now we will elaborate on standardization efforts. One organization is of eminent importance: ISO, which stands short for International Organization for Standardization and is since 1947 the world's largest developer of (voluntary) international standards, which are intended to provide state of the art specifications for products, services and good practices, helping to make industry more efficient and effective. Currently there are more than 19,500 standards available covering almost all aspects of technology and business, from food safety to computers, and agriculture to healthcare (see: www.iso.org).

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.*

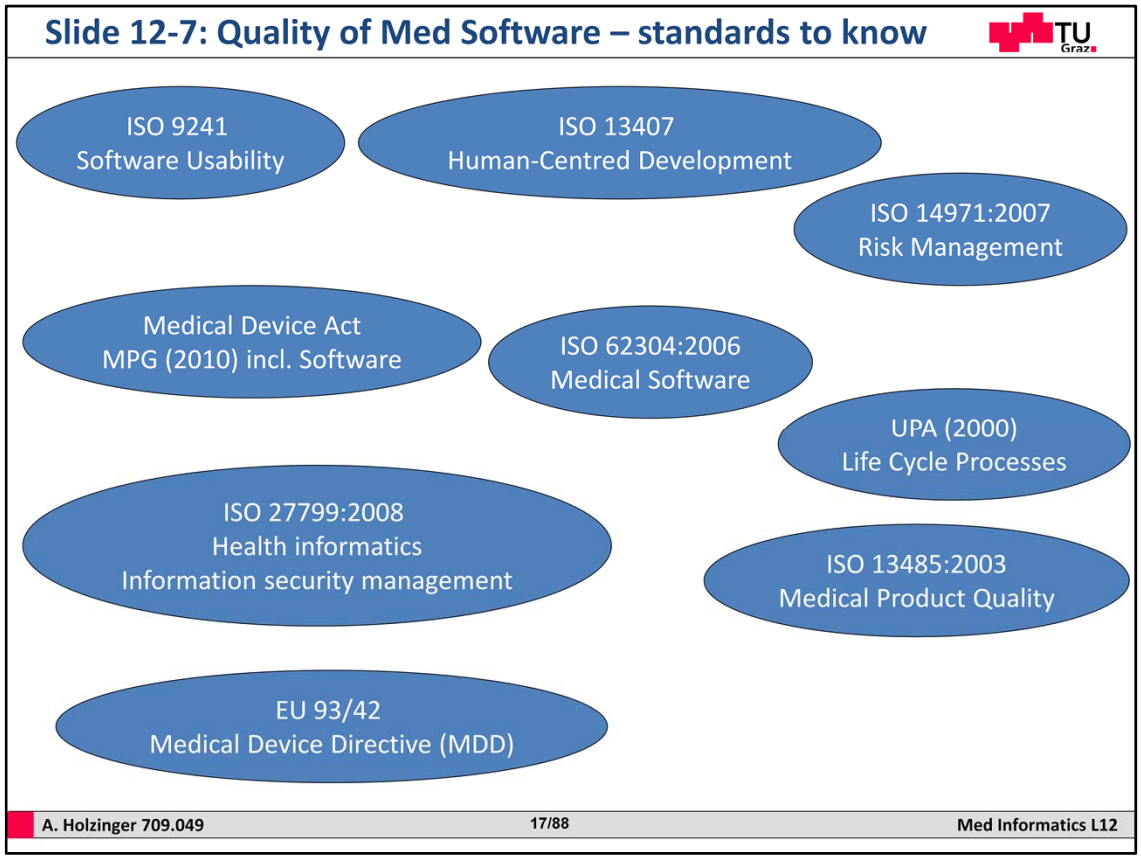
A. Holzinger 709.049 16/88 Med Informatics L12

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 this directive are devices that serve the following purposes (Neuhaus, Polze & Chowdhuryy, 2011): 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 that indicates conformity with the requirements on medical devices of this directive. These requirements are defined in Annex I of the directive and include: 5) Device may not compromise the clinical condition or the safety of patients when used in the intended way; 6) Risks have to be minimized (elimination of risks through security by design, alerts have to warn about dangerous conditions, users have to be informed about residual risks).

Further detailed requirements concern sterility, used materials in manufacturing, influence or emittance of radiation etc.

Devices are classified into risk categories I, IIa, IIb and III depending on the typical duration of use, degree of invasiveness and inherent risk. Category III indicates the highest risk. The requirements for the attainment of a CE mark depend of the risk category the device is classified into. Class III -devices must be approved by the corresponding authority in a EU country prior to market placement and may involve clinical trials (Neuhaus, Polze & Chowdhuryy, 2011).



The International Organization for Standardization (ISO and the International Electro-technical Commission (IEC) provides best practice recommendations on information security risks, management and controls through its ISO/IEC 27000-series standards. The standards cover the fundamental requirements of information management systems; provide 13 guidelines and principles for the implementation of such systems. Among the standards, ISO 27799:2008 and ISO/TR 27809:2007 meant for health informatics and provides guidelines for designing health sector specific information management systems. ISO/IEC 27002 provides control guidelines for patient safety within such systems. ISO/IEC Joint Technical Committee 1 (JTC1) deals with all matters of Information Technology including develop, maintain, promote and facilitate IT standards by enterprises and users concerning the security of IT systems and information.

SO 27799:2008 defines guidelines to support the interpretation and implementation in health informatics of ISO/IEC 27002 and is a companion to that standard.

ISO 27799:2008 specifies a set of detailed controls for managing health information security and provides health information security best practice guidelines. By implementing this International Standard, healthcare organizations and other custodians of health information will be able to ensure a minimum requisite level of security that is appropriate to their organization's circumstances and that will maintain the confidentiality, integrity and availability of personal health information.

ISO 27799:2008 applies to health information in all its aspects; whatever form the information takes (words, numbers, sound recordings, drawings, video, images etc.), whatever means are used to store it (printing or writing on paper or electronic storage) and whatever means are used to transmit it (by hand, fax, over computer networks, mail etc.), as the information must always be appropriately protected.

ISO 13485, published in 2003, represents the requirements for a comprehensive management system for the design and manufacture of medical devices. This standard supersedes earlier documents. ISO 13485 is generally harmonized with ISO 9001. A fundamental difference, however, is that ISO 9001 requires the organization to demonstrate continual improvement, whereas ISO 13485 requires only that they demonstrate the quality system is implemented and maintained.

ISO 14971 is an international standard that is quickly being recognized as one of the best processes to ensure that all aspects of risk management are considered throughout the product lifecycle for medical devices. Compliance to this standard is required to sell medical devices in the European Economic Area, as indicated in the Medical Devices Directive (MDD), which covers most implants, sets conformity assessment procedures depending on the medical device class type, and requires risk analysis to be performed. The use of this standard is also required in Canada and Australia. Within the United States, the standard is recognized by the Food and Drug Administration (FDA) as a way to meet the intent of the Quality System Regulation requirements for the development of safe medical products. ISO 14971 concerns the application of risk management and it is designed to help manufacturers introduce safe medical devices into the healthcare market. The manufacturer is responsible for identifying and controlling not only the risks associated with their medical device, but evaluating interactions with other devices. The standard also allows for other healthcare manufacturing organizations to use the process and obtain certification. This might include human tissue, animal care products, pharmaceutical manufacturers, etc., who may choose to use this standard. (Catelani et al., 2011).

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

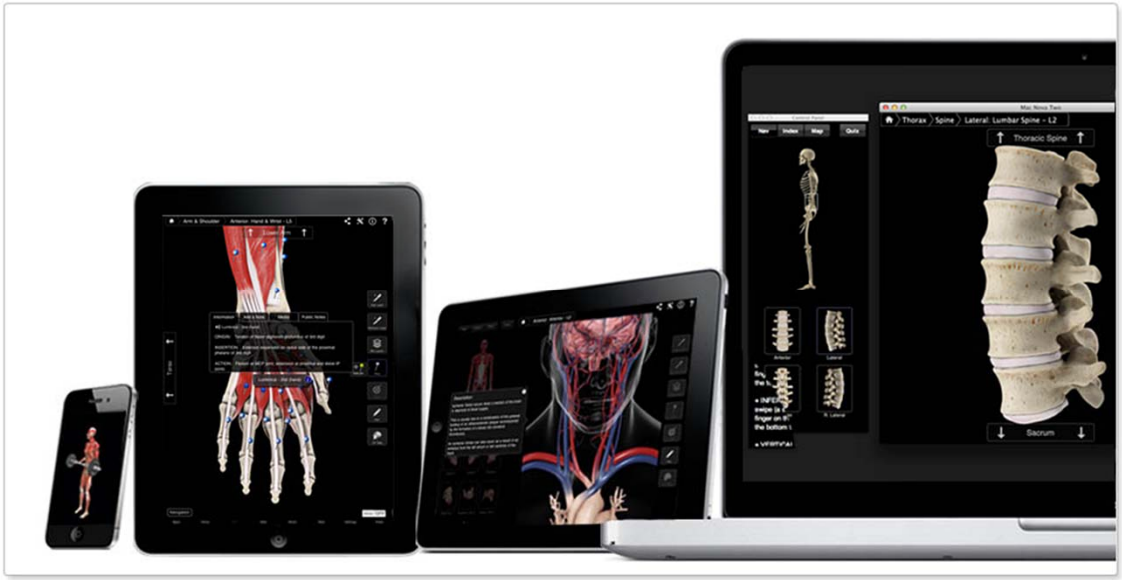
Änderung des Medizinproduktegesetzes

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

A. Holzinger 709.049 19/88 Med Informatics L12

Slide 12-9 Medical Product Law and mobile Apps
On September 25, 2013, the FDA (Food and Drug Administration, see: http://www.fda.gov/medicaldevices/productsandmedicalprocedures/connectedhealth/mobilemedicalapplications) released a (non-binding) document on Mobile Medical Applications recommendations.
The widespread adoption of mobile computing in medicine and in particular the success of mobile Applications (Apps) is opening new and innovative ways to improve medicine, health and health care delivery (Peischl, Ferk & Holzinger, 2013), (Breitwieser et al., 2013), (Novak et al., 2012), (Holzinger et al., 2011).
Apps can also help to manage personal health and wellness and promote healthy living (Alagoez et al., 2010), (Holzinger et al., 2010). According to industry estimates, 500 million smartphone users worldwide will be using a health care application by 2015, and by 2018, 50 percent of the more than 3.4 billion smartphone and tablet users will have downloaded mobile health applications (http://www.research2guidance.com/500m-people-will-be-using-healthcare-mobile-applications-in-2015). These users include health care professionals, consumers, and patients.
The FDA encourages the development of mobile medical apps that improve health care and provide consumers and health care professionals with valuable health information. The FDA also has a public health responsibility to oversee the safety and effectiveness of medical devices – including mobile medical apps, for this purpose the FDA issued the Mobile Medical Applications Guide, which explains the oversight of mobile medical apps as devices and our focus only on the apps that present a greater risk to patients if they don't work as intended and on apps that cause smartphones or other mobile platforms to impact the functionality or performance of traditional medical devices.

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.

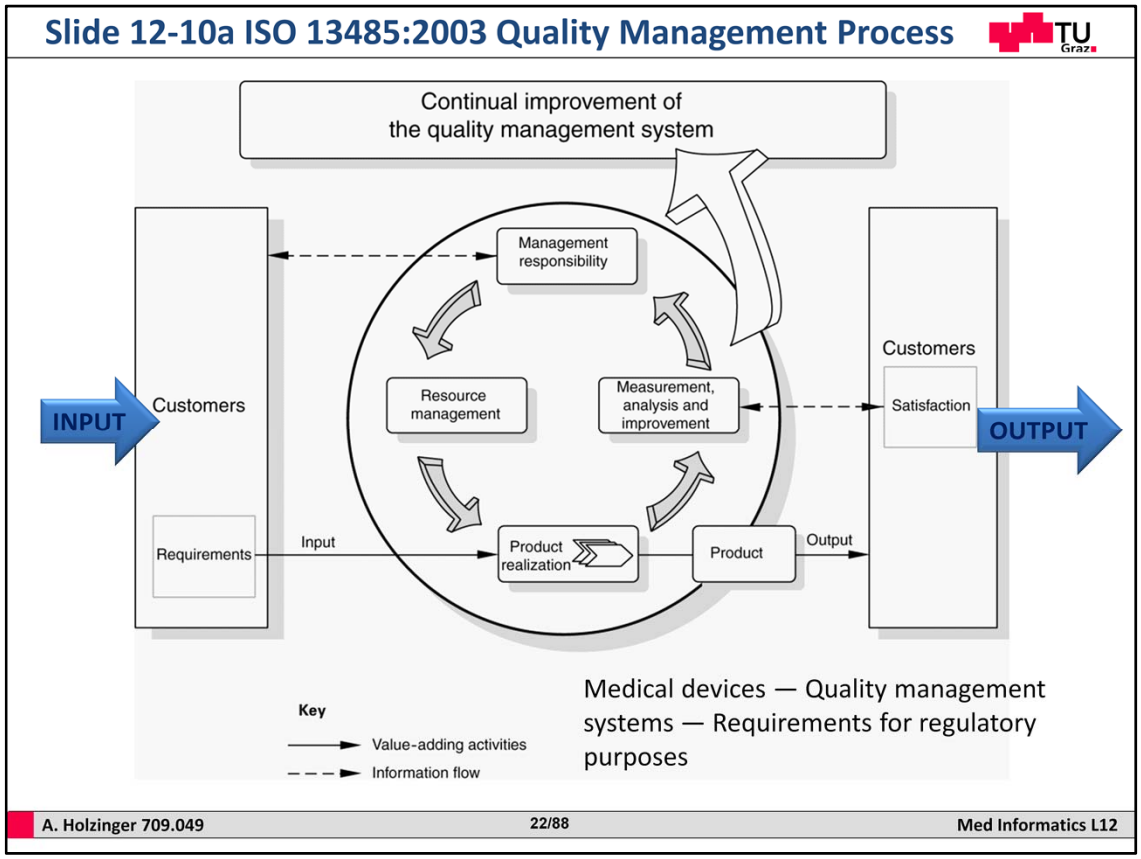
A. Holzinger 709.049 | 21/88 | Med Informatics L12

In the traditionally developed university system, there has been a cultural gap between the classical natural basic sciences (e.g., chemistry, biology, physics) and applied fields such as engineering or clinical medicine, the latter many believe to be more an art than a science (Kuhn et al., 2008). If we look what both sides have in common, it is obvious that information and quality are in both areas considered as important. Consequently, modern information management can bridge the hiatus theoreticus, the gap between (scientific) knowledge and its application (Simonic & Holzinger, 2010).

References:

Simonic, K.-M. & Holzinger, A. 2010. Zur Bedeutung von Information in der Medizin. OCG Journal, 35, (1), 8.

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



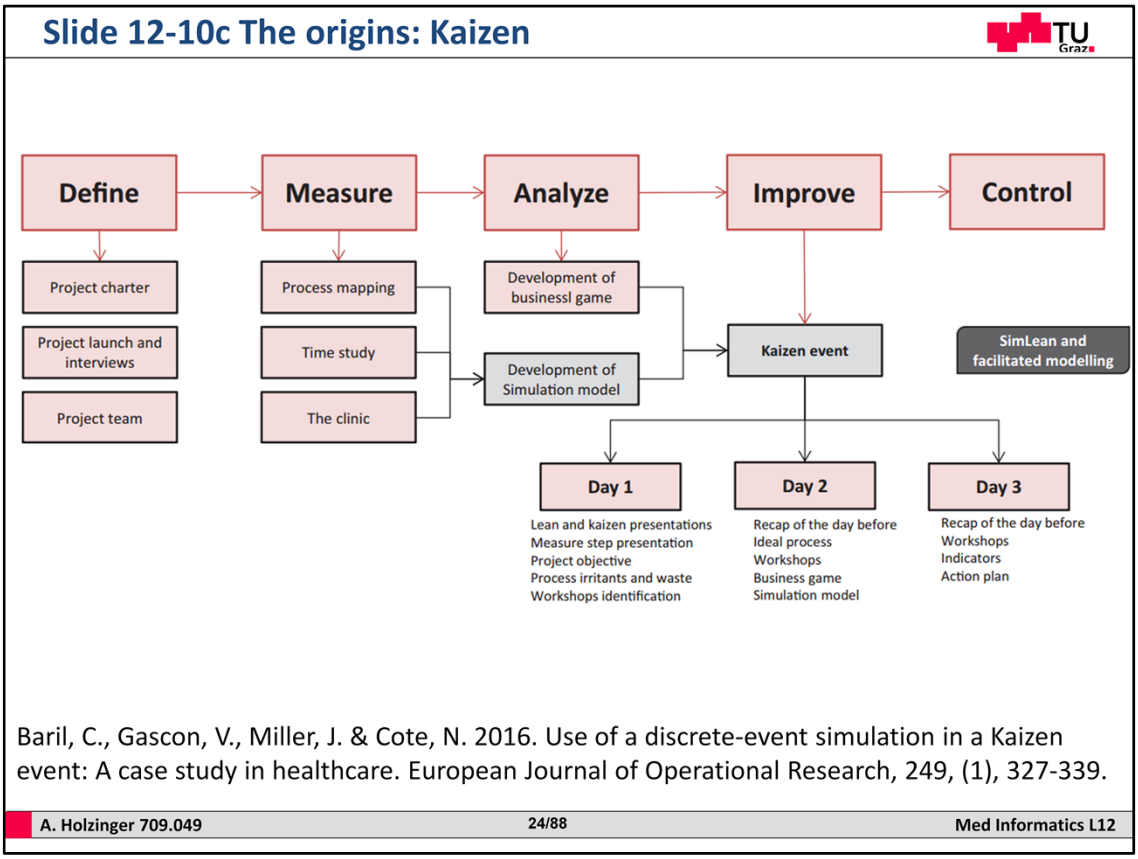
Slide 12-10: ISO 13485:2003 Quality Management Process Cycle
ISO 13485:2003 represents the requirements for a comprehensive management system for the design and manufacture of medical devices. More specific it describes the requirements for a quality management system where an organization (regardless of size or type) needs to demonstrate its ability to provide medical devices and related services that consistently meet customer requirements and regulatory requirements. In this slide we see the main idea behind it: The Quality Management Process Cycle: The customers (aka end-users) specify the requirements as active input for the product realization. Within the cycle we have a consequent iteration, forth back checking if the requirements are met, similar to the PDCA cycle (Holzinger, 2011) – see next slide. The ideal output is in satisfactory addressing all end user requirements.

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

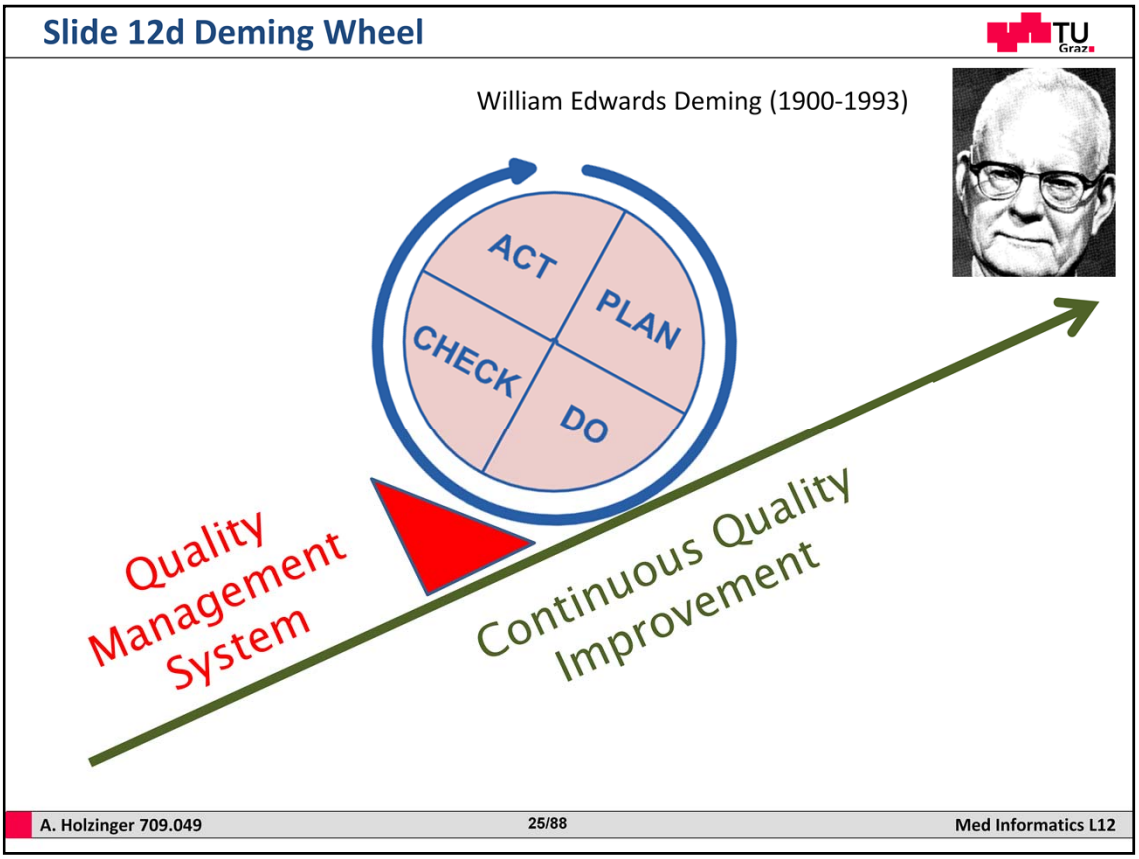
Masaaki, I. 1986. Kaizen: The Key to Japan's Competitive Success. New York: Random House.

Kai zen is Japanese and means change good or change for the better = continuous improvement



This is just to show you that this is not an antique approach – here a very recent study by a group from the University of Ontario: they studied how a business game can be used jointly with discrete event simulation to test scenarios defined by team members during a Kaizen event. The aim was to allow a rapid and successful implementation of the solutions developed during the Kaizen. It has been used to improve patients' trajectory in an outpatient hematology–oncology clinic. Patient delays before receiving their treatment were reduced by 74 percent after 19 weeks.

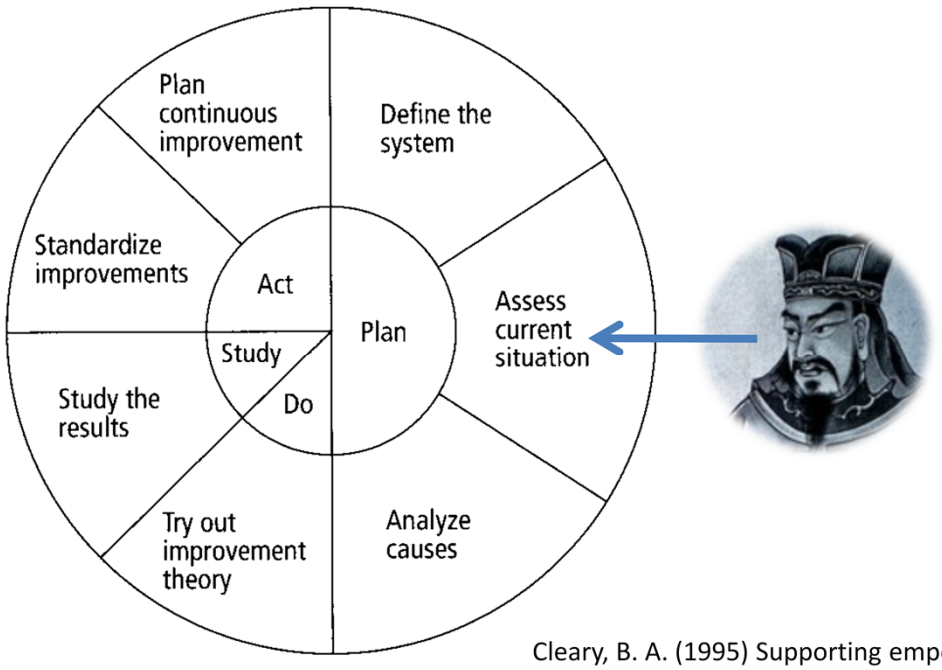
The roots of this go back to the implementation of Kaizen by



The famous words of Deming: “If you do not know how to ask the right question, you discover nothing” and if you cannot measure it, you cannot manage it.
and Steve Jobs said: “customers do not measure you on how hard you tried and how much time you spent – they measure you on what you deliver!”

PLAN
Establish the objectives and processes necessary to deliver results in accordance with the expected output (the target or goals). By establishing output expectations, the completeness and accuracy of the spec is also a part of the targeted improvement. When possible start on a small scale to test possible effects.
DO
Implement the plan, execute the process, make the product. Collect data for charting and analysis in the following "CHECK" and "ACT" steps.
CHECK
Study the actual results (measured and collected in "DO" above) and compare against the expected results (targets or goals from the "PLAN") to ascertain any differences. Look for deviation in implementation from the plan and also look for the appropriateness and completeness of the plan to enable the execution, i.e., "Do". Charting data can make this much easier to see trends over several PDCA cycles and in order to convert the collected data into information. Information is what you need for the next step "ACT".
ACT
If the CHECK shows that the PLAN that was implemented in DO is an improvement to the prior standard (baseline), then that becomes the new standard (baseline) for how the organization should ACT going forward (new standards are enACTed). If the CHECK shows that the PLAN that was implemented in DO is not an improvement, then the existing standard (baseline) will remain in place. In either case, if the CHECK showed something different than expected (whether better or worse), then there is some more learning to be done... and that will suggest potential future PDCA cycles. Note that some who teach PDCA assert that the ACT involves making adjustments or corrective actions... but generally it would be counter to PDCA thinking to propose and decide upon alternative changes without using a proper PLAN phase, or to make them the new standard (baseline) without going through DO and CHECK steps.

Slide 12-11: Quality Improvement Cycle

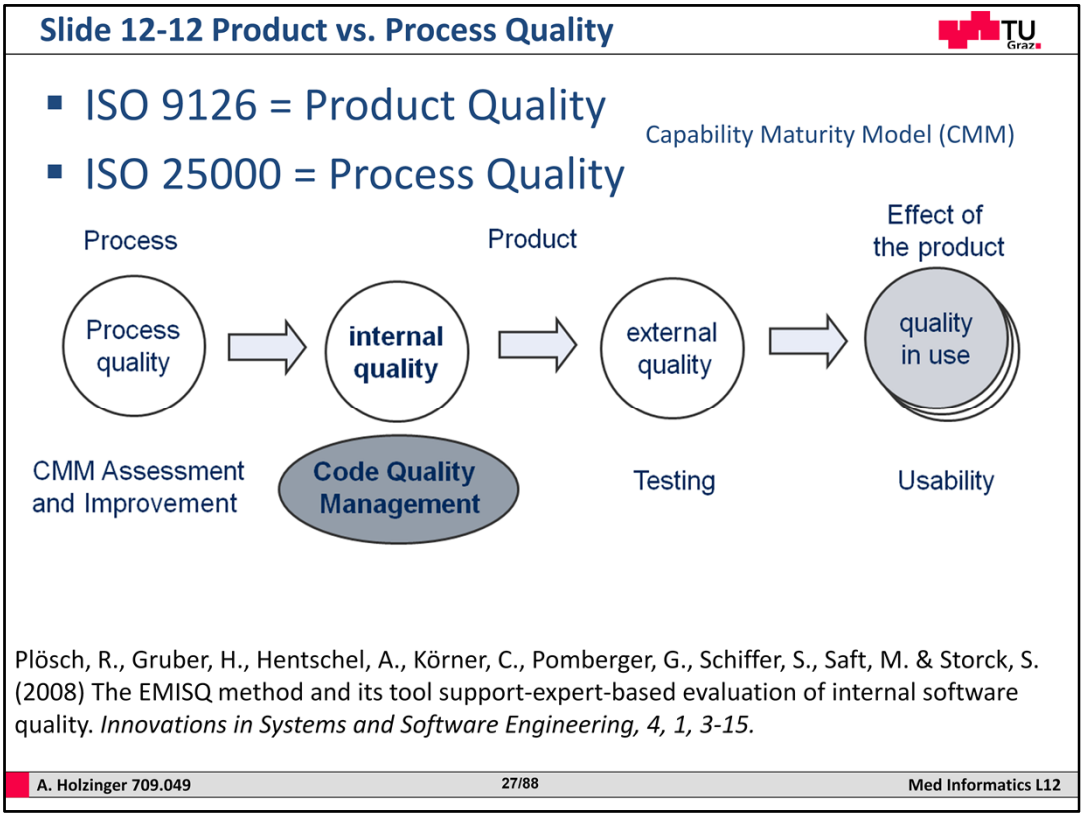


Cleary, B. A. (1995) Supporting empowerment with Deming's PDSA cycle. *Empowerment in Organizations, 3, 2, 34-39.*

A. Holzinger 709.049 26/88 Med Informatics L12

Sun Tzu (544-496 BC) "If you know the enemy and know yourself, you need not fear the result of a hundred battles. If you know yourself but not the enemy, for every victory gained you will also suffer a defeat. If you know neither the enemy nor yourself, you will succumb in every battle."

Slide 12-11: Quality Improvement Cycle
The quality improvement cycle is based on the original PDCA-Cycle aka Deming Cycle (Deming, 1994) – Plan, Do, Study (the results), Act (incorporate your improvements). In this slide it is extended to a seven-step improvement process, which applies to any organization. In hospitals this approach brought for example enormous reductions of waste of supplies (Cleary, 1995); the steps include:
1) Defining the system
2) Assessing the current situation
3) Analyzing causes
4) Applying an improvement process
5) Studying the results
6) Planning continuous improvement: No improvement process is ever finished!
7) Standardize the improvements!
This process is widely adopted in the medical area (Cleary, 1995).

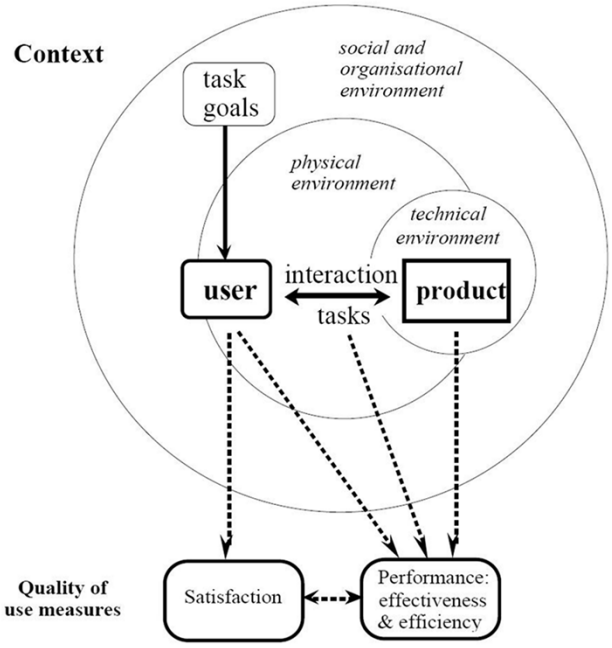


Slide 12-12: Product Quality versus Process Quality
Before we concentrate on the Software Quality Model, we emphasize again the difference between product quality (which is defined in ISO 9126) and process quality (which is defined in ISO 25000) and the important insight that both of them are important for the goal: quality in use (see next slide).

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.*

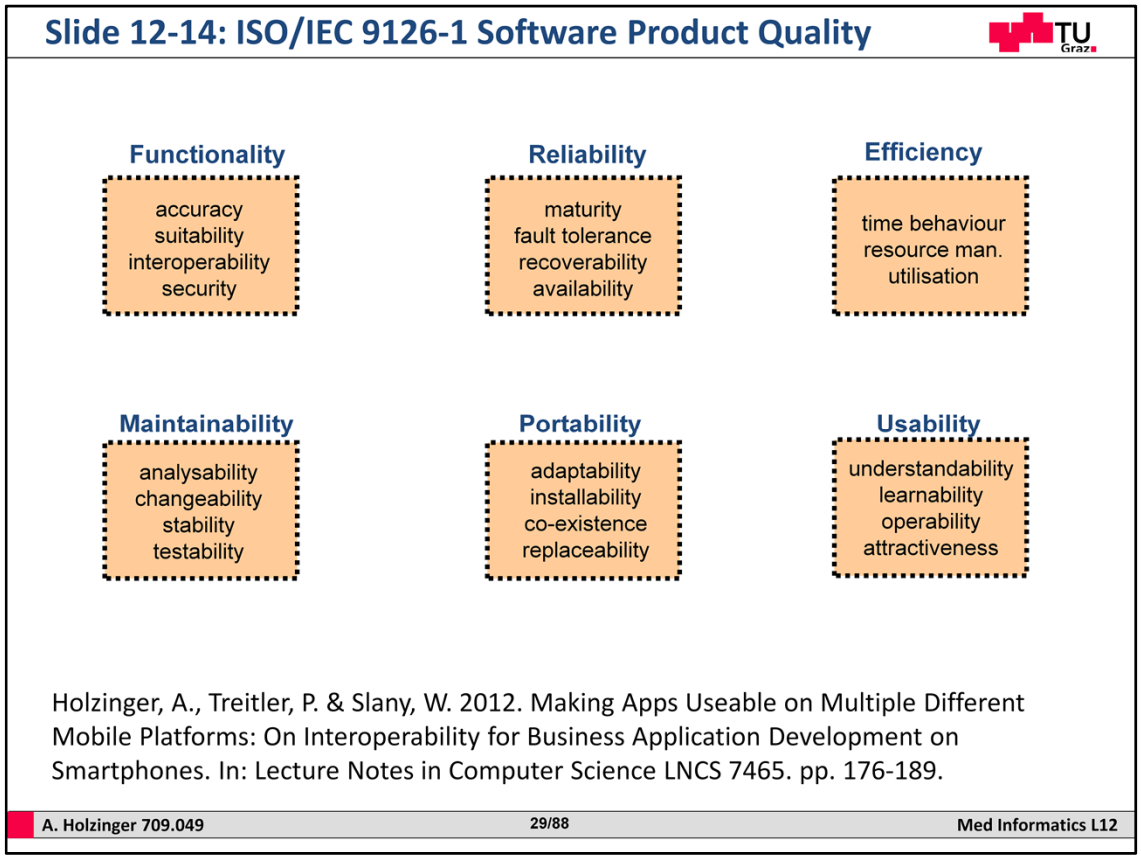
Bevan, N. (1995) Measuring Usability as Quality of Use. *Software Quality Journal, 4,* ***2, 115-130.***



A. Holzinger 709.049 28/88 Med Informatics L12

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

The important insight which we shall always consider is that the quality in use is the goal and that the quality of use measure is "usability" (Bevan, 1995), (Bevan, 1997), (Bevan, 2009), (Holzinger et al., 2009) and this is always taking place within a context wherein the user constantly interacts with the product. Software in that sense is also a product.



Slide 12-14: ISO/IEC 9126-1 Software Product Quality
Usability is important but only a small part within the whole software product quality life cycle. ISO 9126-1 defines six large areas, each containing a set of important issues:
1) Functionality: accuracy, suitability, interoperability, security;
2) Reliability: maturity, fault tolerance, recoverability, availability;
3) Efficiency: time behavior (especially critical in the clinical domain!), utilization
4) Maintainability: analyzability, changeability, stability, testability;
5) Portability: adaptability, installability, co-existence, replaceability;
6) Usability: understandability, learnability, operability, attractiveness;
Let us look closer on issue Nr. 5: Portability: this is particularly important with Apps: Making Apps useable on different platforms: the shipment of smartphones exceeded that of personal computers in 2011. However, the screen sizes and display resolutions of different devices vary to a large degree, along with different aspect ratios and the complexity of mobile tasks. These obstacles are a major challenge for software developers, especially when they try to reach the largest possible audience and develop for multiple mobile platforms or device types. On the other side, the end users' expectations regarding the usability of the applications are increasing. Consequently, for a successful mobile application the user interface needs to be well-designed, thus justifying research to overcome these obstacles. In this paper, we report on experiences during an industrial project on building user interfaces for database access to a business enterprise information system for professionals in the field.
Holzinger, Treitler & Slany (2012) discuss a systematic analysis of standards and conventions for design of user interfaces for various mobile platforms, as well as scaling methods operational on different physical screen sizes.

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.*

A. Holzinger 709.049 30/88 Med Informatics L12

Slide 12-15: Remember Medical workflows are highly complex …
One of the basic lessons from HCI is, that usability must be considered before prototyping takes place. There are techniques (such as usability context analysis) intended to facilitate such early focus and commitment. When usability inspection, or testing, is first carried out at the end of the design cycle, changes to the interface can be costly and difficult to implement, which in turn leads to usability recommendations. These are often ignored by developers who feel, “We don’t have usability problems.” The earlier critical design flaws are detected, the more likely they can be corrected. Thus, user interface design should more properly be called user interface development, analogous to software development, since design usually focuses on the synthesis stages, and user interface components include metaphors, mental models, navigation, interaction, appearance, and usability (Holzinger et al., 2005b).

Slide 12-16: Comparison of Usability Engineering Methods

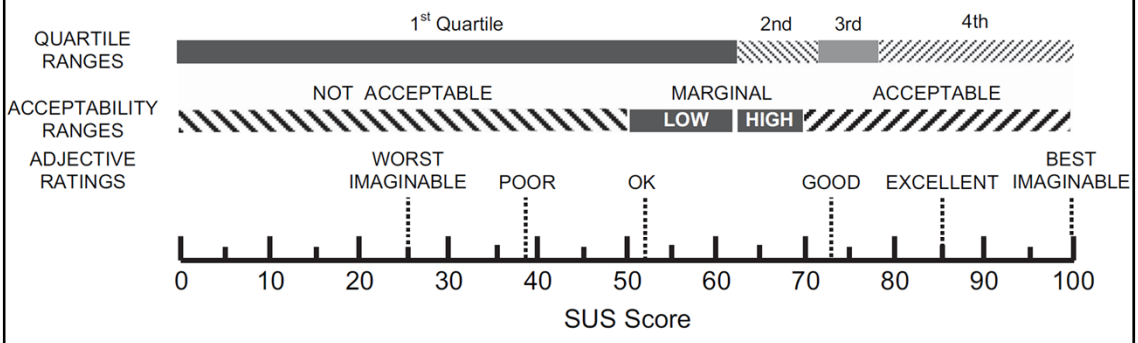
	Inspection Methods			Test Methods		
	Heuristic Evaluation	Cognitive Walkthrough	Action Analysis	Thinking Aloud	Field Observation	Questionnaires
Applicably in Phase	all	all	design	design	final testing	all
Required Time	low	medium	high	high	medium	low
Needed Users	none	none	none	3+	20+	30+
Required Evaluators	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.*

A. Holzinger 709.049 31/88 Med Informatics L12

Generally, we can determine between two types of usability engineering methods: Inspection vs. Test (Holzinger, 2005).
Inspection methods are a set of methods for identifying usability problems and improving the usability of an interface design by checking it against established standards. These methods include heuristic evaluation, cognitive walkthroughs, and action analysis. No end users are needed, these methods are performed by experts.
Testing with (real) end users is the most fundamental usability method and is in some sense indispensable. It provides direct information about how people use our systems and their exact problems with a specific interface. There are several methods for testing usability, the most common being thinking aloud, field observation, and questionnaires.

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.*

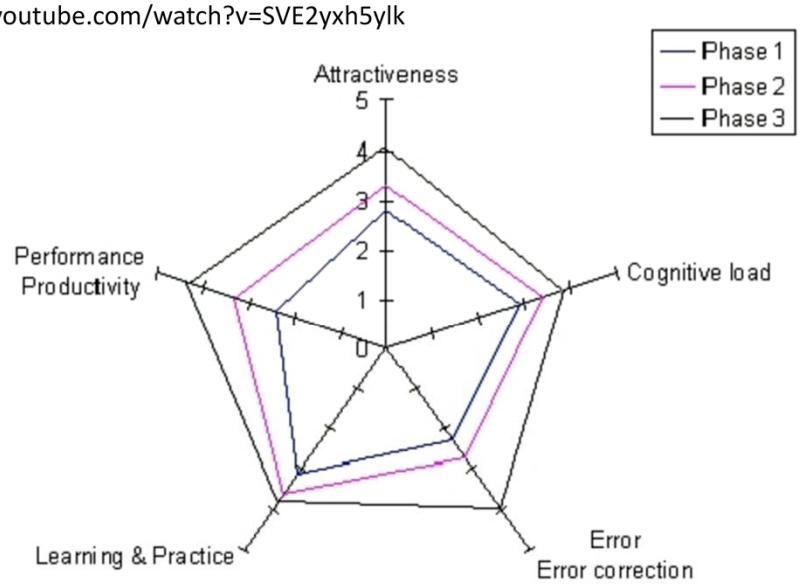
A. Holzinger 709.049 32/88 Med Informatics L12

Slide 12-17: The System Usability Scale (SUS)
A rapid evaluation tool is the System Usability Scale (SUS). This 10-item scale was developed by (Brooke, 1996) as a "quick and dirty" survey scale that would allow the usability practitioner to quickly and easily assess the usability of a given product or service. Although there are a number of other excellent alternatives the SUS has several attributes that make it a good choice for general usability practitioners. The main advantage is, that the survey provides a single score on a scale that is easily understood by the wide range of people (from project managers to computer programmers) who are typically involved in the development of products and services and who may have little or no experience in human factors and usability (Holzinger, 2010).

http://www.measuringusability.com/blog/10-things-SUS.php
http://www.measuringusability.com/products/SUSpack

Slide 12-18: Software Usability Measurement Inventory SUMI

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

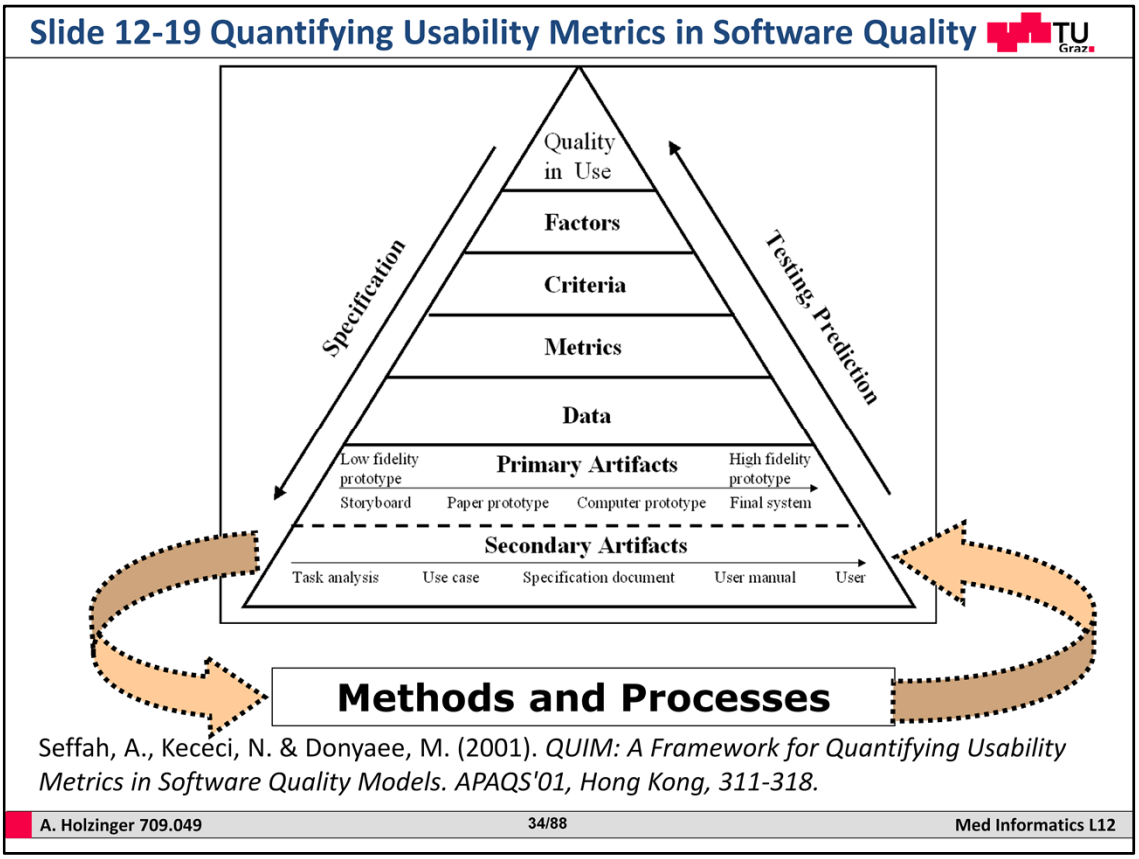


Kosec, P., Debevc, M. & Holzinger, A. 2009. Towards Equal Opportunities in Computer Engineering Education: Design, Development and Evaluation of Video-based e-Lectures. International Journal of Engineering Education (IJEE), 25, (4), 763-771.

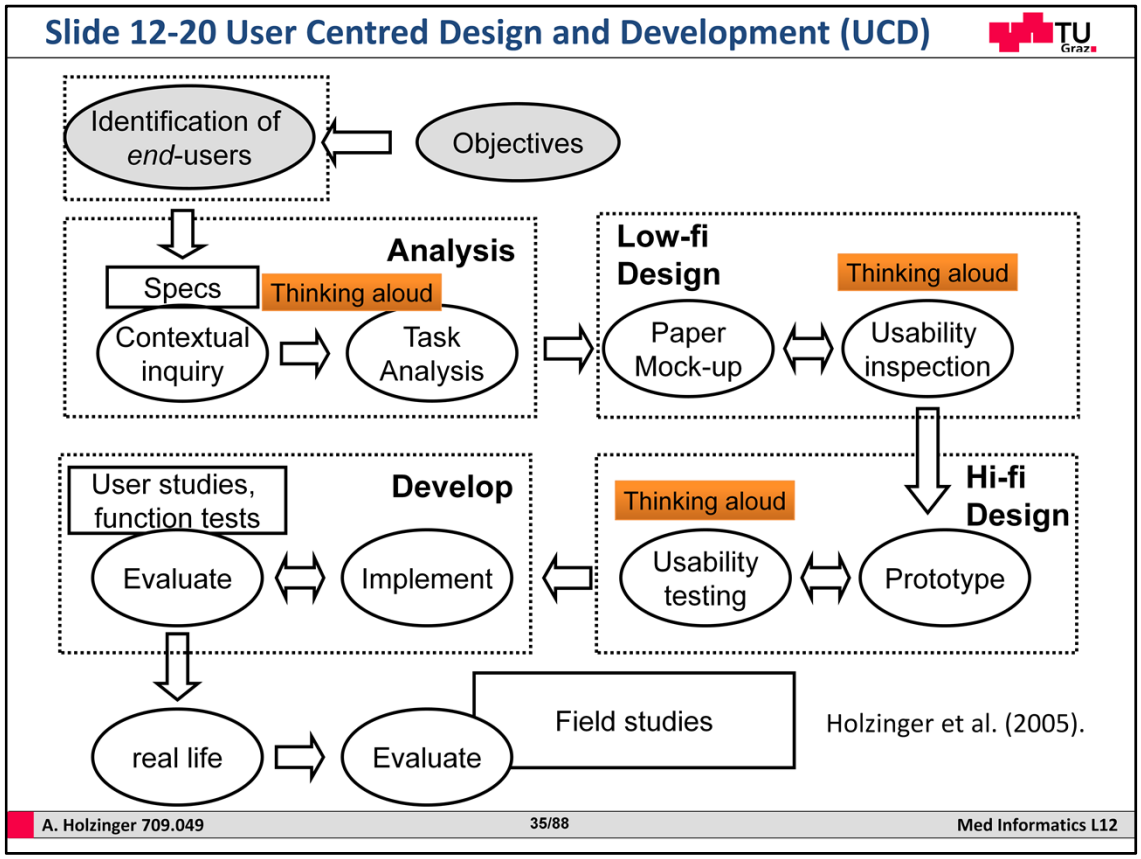
A. Holzinger 709.049 33/88 Med Informatics L12

Slide 12-18: The Software Usability Measurement Inventory (SUMI)
According to (Kirakowski & Corbett, 1993) the assessment of the usability of a computer system should involve measuring not only aspects of users' performance, but also how users subjectively feel about the system. For this purpose the Software Usability Measurement Inventory (SUMI) has been designed in particular to investigate users' perceptions of the quality of software systems. SUMI provides a global usability measure, along with five subscale measures and a high level problem diagnosis. There are large samples available, which can be used as benchmarks tested either against generic usability profiles, or against the usability profile of another system. An sample application of a SUMI evaluation can be found in (Kosec, Debevc & Holzinger, 2009) and a good discussion in (Cavallin, Martin & Heylighen, 2007) and a good source is available here:
http://sumi.ucc.ie/sumipapp.html

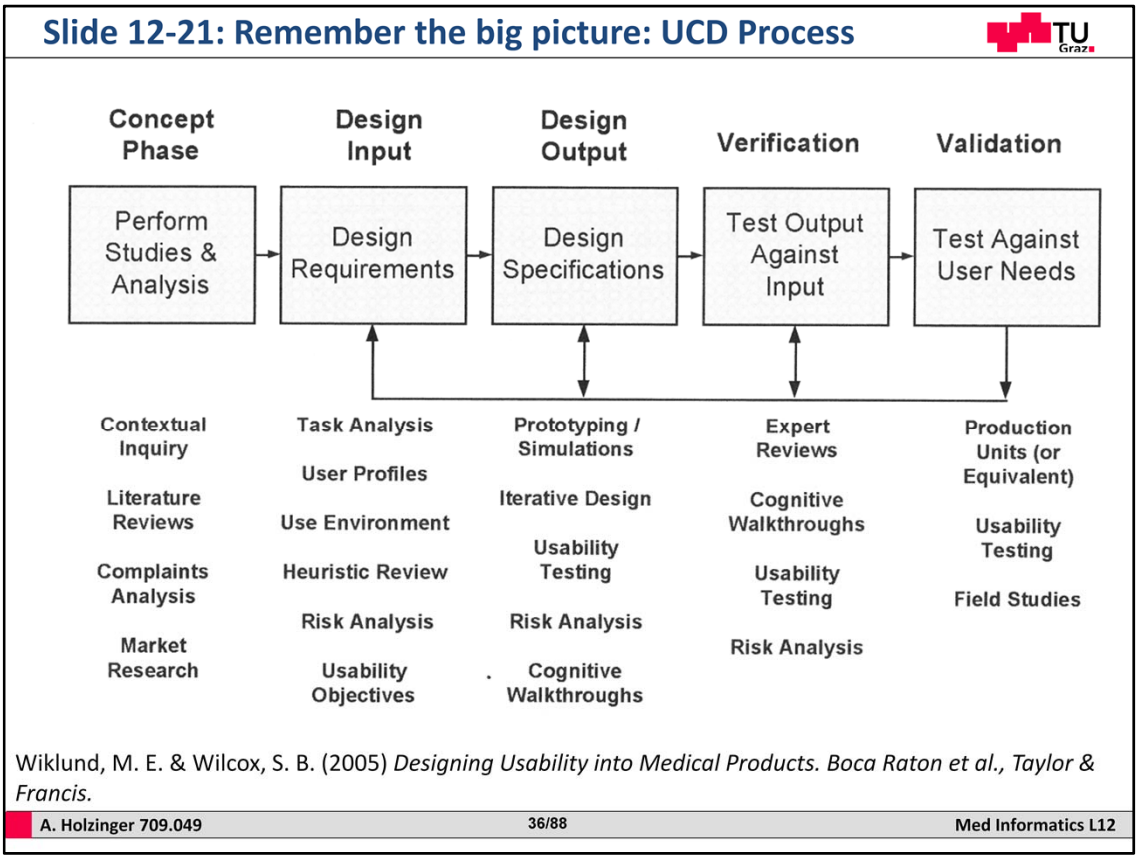
A funny video can be found here:
http://www.youtube.com/watch?v=SVE2yxh5ylk



Slide 12-19 Quantifying Usability Metrics in Software Quality
In this slide we see QUIM: A Framework for Quantifying Usability Metrics in Software Quality Models, which is a hierarchical model similar to typical software engineering models (e.g. Boehm model, McCall model, IEEE 1061, ISO 9126, etc.). The difference is that, it distinguishes four levels called factors, criteria, metrics and data – as can be seen in the slide. The relationship between these layers is an n–m relationship. Factors include effectiveness, efficiency, satisfaction, productivity, safety, internationability (globality); the criteria include attractiveness, consistency, minimal action, minimal memory load, completeness; the metrics include task concordance and visual coherence (Seffah, Kececi & Donyaee, 2001), (Seffah et al., 2006), (Holzinger et al., 2008).



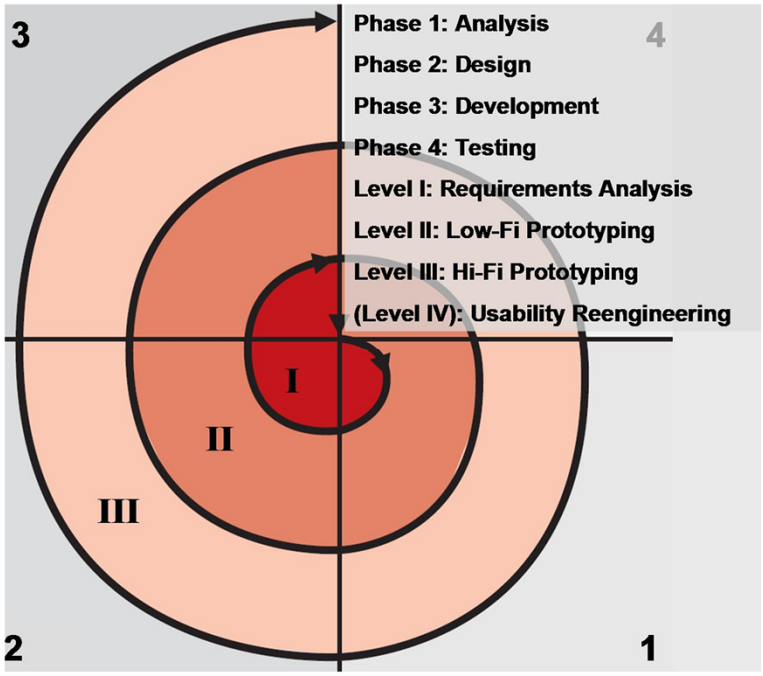
In contrast to traditional ego-centric design , user centred design and development focuses on the needs, demands, and requirements of the end user.
Note: In software engineering the design is more the "thinking" and problem solving, e.g. the solution of a problem in terms of algorithms and concepts; whereas development includes the implementation of the design. Usually, an engineer performs both: design and development. The emphasis in this model is on the end user. Meanwhile the ISO 13407 standard defines a Human-centred design process, which defines a general process model (similar to the "big picture" in the next slide) but does not define specific methods. In this slide we see a process model, which has been proved in many projects and where for example thinking aloud as a main low-cost method can be applied from the very early stages of the development cycle. The most important step is to identify the end users at the very beginning, then to specify the context of use, create low-fi design solutions, because they can be redesigned rapidly and with low-cost (Holzinger, 2002), (Holzinger, 2003), (Holzinger, Wascher & Steinmann, 2003).



Slide 12-21: Remember the big picture: UCD Process
In this slide we see the "big picture": The UCD process as strategy of the whole development process (Wiklund & Wilcox, 2005) includes the concept phase with contextual inquiries and market research; the design requirement phase with task analysis and user profiling, the design specifications phase with the first prototypes a the verification phase with usability testing and finally the validation phase with field studies and evaluation (we come to this in the next slides).

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.*

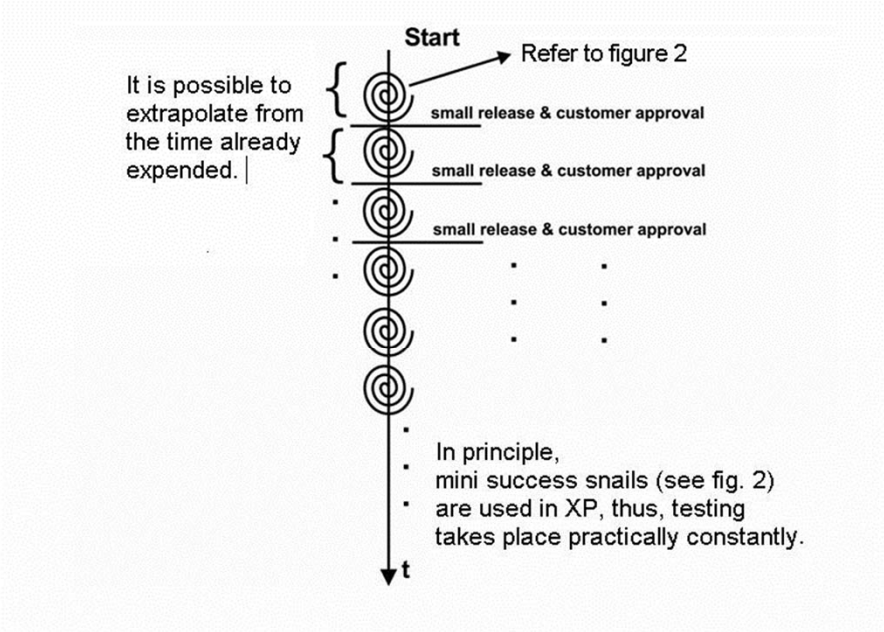


A. Holzinger 709.049 37/88 Med Informatics L12

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

The success of Extreme Programming (XP) is based, among other things, on an optimal communication in teams of 6-12 persons, simplicity, frequent releases and a reaction to changing demands (Beck, 1999). Most of all, the customer (hence: end user) is integrated into the development process, with constant feedback. This is very similar to Usability Engineering (UE) which follows a spiral four phase procedure model (analysis, draft, development, test) and a three step (paper mock-up, prototype, final product) production model. In comparison, these phases are extremely shortened in XP; also the ideal team size in UE User-Centered Development is 4-6 people, including the end-user. The two development approaches have different goals but, at the same time, employ similar methods to achieve them. It seems obvious that there must be synergy in combining them. The authors present ideas in how to combine them in an even more powerful development method called Extreme Usability (XU) (Holzinger et al., 2005a), (Holzinger & Slany, 2006), (Hussain, Slany & Holzinger, 2009a), (Hussain, Slany & Holzinger, 2009b).

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

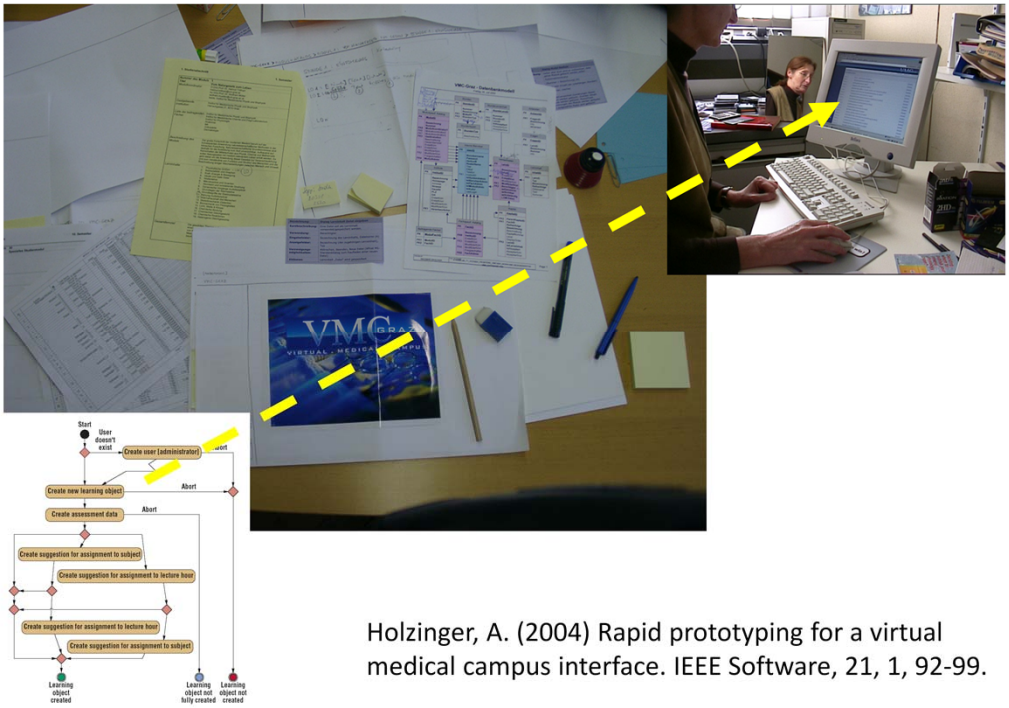


Holzinger, A. & Slany, W. (2006) XP + UE -> XU Praktische Erfahrungen mit eXtreme Usability. *Informatik Spektrum, 29, 2, 91-97.*

A. Holzinger 709.049 38/88 Med Informatics L12

By making the UCD spirals (Slide 12-22) as small as possible we achieve a series of advantages:
In XP, this danger of dissipating one's energies in details (engineers are particularly susceptible to "featuritis" (Buschmann, 2010)) and the client's (end users who are also often overstressed) becomes caught up in the detail is consciously controlled by applying short iterations, frequent re-planning and focusing on simple design.
Note: Simple things first – they may be the most important ones!
This enables the client to get a realistic feeling of what can be achieved by the team, if the team implements only what he requested, and what needs to be pushed back to later versions in order to achieve the core functionality needed for the economic success of the project.
In particular, the well-known danger of "featuritis" is harnessed by the conscious decision to avoid thinking about what could happen later and could become meaningful, while being prepared to make extensive adjustments and changes at a later date. Extreme Usability (XU) could become such that all the best practices of UE are kept in the XP process during the planning games, with a restriction of the usability aspects in the next iteration and the equal treatment of Usability and Functionality.
The advantage would be that, with the XP process, the adjustment and gradual improvement until the end of the project is explicitly built into the process, which is very helpful for UE. However, UE can improve the XP development method by focusing on the important aspects of the usability and employing the entire development team to make the customer continually aware of these aspects (by daily inquiry, discussion and testing); also the developers minds will be focused on the most important usability aspects, when at least one developer in the team possesses previous knowledge about UE and by implementing pair-programming, including the complete and frequent mixing of the pairs as well as passing on the On-Site Customer XP principle. Obviously, UE experience for all developers is an advantage in every project (Holzinger et al., 2005).

Slide 12-24 Rapid Prototyping – Paper Mock-ups



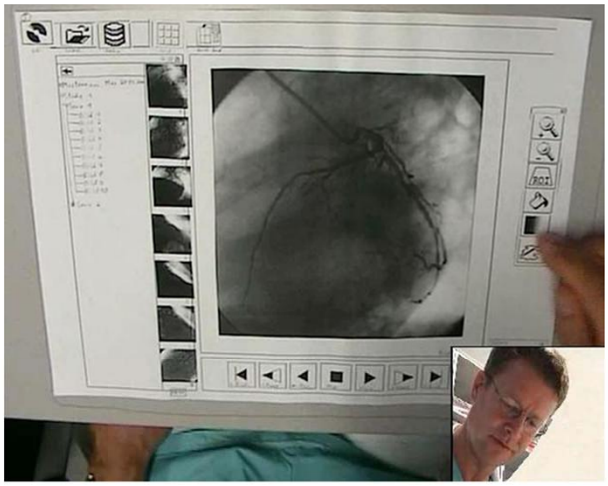
Holzinger, A. (2004) Rapid prototyping for a virtual medical campus interface. IEEE Software, 21, 1, 92-99.

A. Holzinger 709.049 39/88 Med Informatics L12

A very helpful Lo-fidelity prototyping method is to use paper mock-ups for rapid prototyping. Using common office supplies (markers, index cards, scissors and transparency film) the engineer can quickly sketch screen contents and each interactive element of the interface (menus, messages) on a separate piece of paper. The paper mock-up is not necessary to be very neat: it may contain hand written text, crooked lines and last minute corrections. It is, however, good enough to show what the screens would look like and provides a good basis for "playing out" some workflows. One developer plays the role of the "computer," simulating the behavior of the software by manipulating the pieces of paper. It is important to ask the end users to perform realistic tasks with the prototype, e.g. "... you are a teacher, set up a theme and create some hours in the catalogue ..." Furthermore, it is important NOT to ask the users for their opinions of the interface – telling them that this is experimental is enough. After every UCD session the team can discuss what they had seen and immediately execute changes to the paper prototype (Holzinger, 2004).

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.

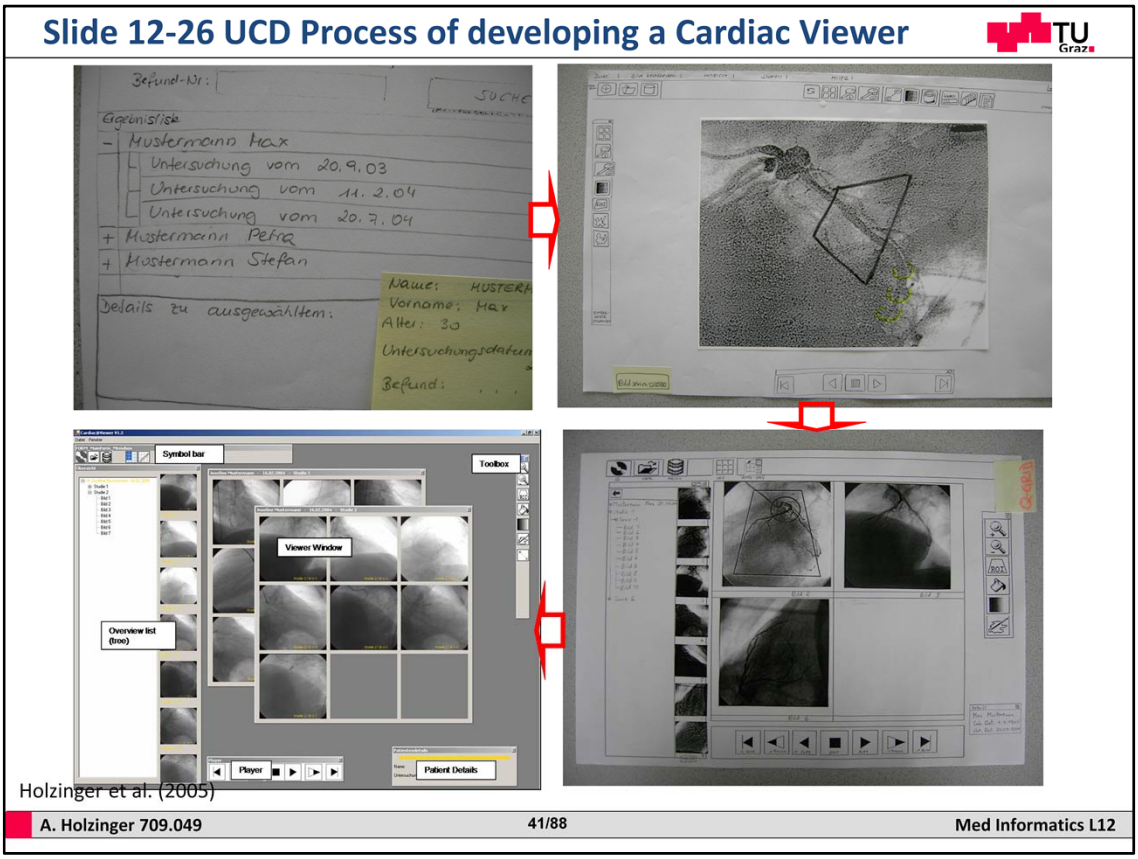


Brown, S. & Holzinger, A. (2008) Low cost prototyping: Part 1, or how to produce better ideas faster by getting user reactions early and often. In: Abuelmaatti, O. & England, D. (Eds.) *Proceedings of HCI 2008. Liverpool: John Moores University (UK), British Computer Society, 213–214.*

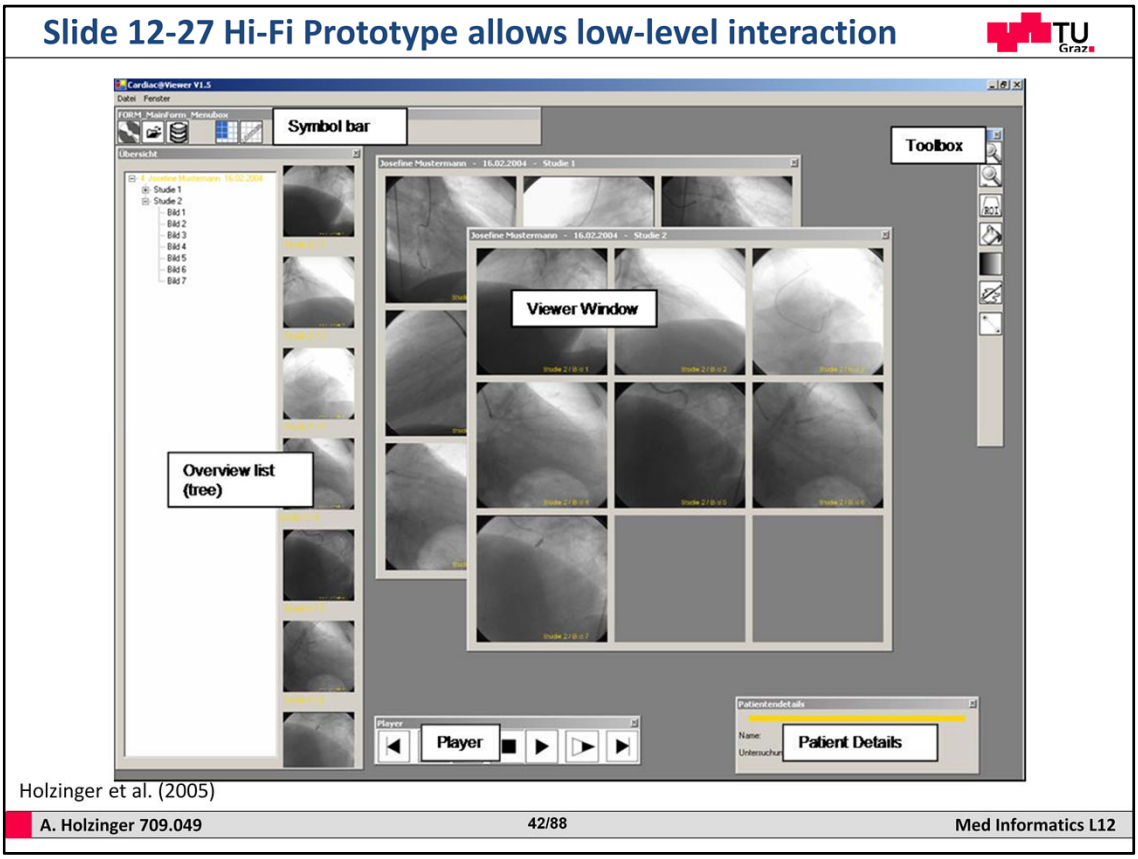
Holzinger, A. & Brown, S. (2008) Low cost prototyping: Part 2, or how to apply the thinking-aloud method efficiently. In: Abuelmaatti, O. & England, D. (Eds.) *Proceedings of HCI 2008. Liverpool: John Moores University (UK), British Computer Society, 217–218.*

A. Holzinger 709.049 40/88 Med Informatics L12

From the spectrum of methods in Usability Engineering (review →Slide 12-16), one method particularly stands out due to its practical realizability: Thinking aloud (THA). This method originates from early psychological problem solving research (Duncker, 1945) and permits insight into the mental processes: The test person (end user) receives a completely defined set of tasks and is asked to express, out loud, all – also fleeting – ideas and thoughts during the execution of this task. It is advantageous to record this procedure with a video camera because it is then possible to draw conclusions on the work habits from both the verbal and the facial expression and the gestures of the test person, in particular, it is possible to judge their subjective impressions and feelings. The behaviour patterns recorded on the video tape and/or the log file analyses usually make it possible to identify where the test person has problems and how, and why, they take certain actions. Additionally, with a behaviour observation software (for example INTERACT from the company Mangold Munich, Germany), the video material can also be compared to that of other users in order to find particular behaviour patterns. According to (Nisbett & Wilson, 1977) 3 to 5 end users are sufficient to obtain valid statements; however, for scientific studies, it is sometimes necessary to increase the number of test subjects. The principle of breaking off the tests when no further increase in knowledge is effected, has proved satisfactory (Brown & Holzinger, 2008);(Holzinger & Brown, 2008).

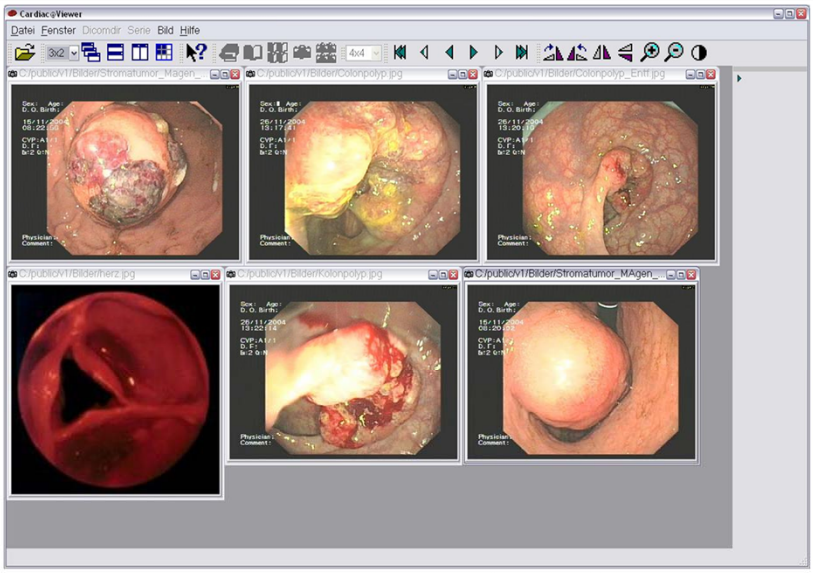


This practical example from a development project (Holzinger et al., 2005b) shows the four levels of the UCD process:
Level 1: Requirements Analysis: The first goal is to provide specifications of the tasks that the end users must perform in order to support problem solving. The envisioned system shall be discussed with all the people involved (not the CEO, who never will do any work with it ;-) in the medical example here it includes the cardiologists, who expressed their demands of what the system must be able to do – which functionalities it should provide and how these should work. During the requirement analysis, which was made with the help of video recordings of clinicians in real work situations, a verbal description of the system emerged.
Level 2: Low-Fi Prototyping (Paper Mock-up): In this project, at first, screen designs and dialogues were sketched on paper. Then a paper mock-up, which can be adjusted whilst working with the cardiologists, was created. The use of paper mock-ups provides a first usability feedback with minimum effort and maximum results.
Level 3: Hi-Fi Prototyping: Further, a working prototype, for studying the interaction of the end users, was created. During this phase the programmers were able to concentrate on the hi-fi prototype and adapt the choice of software tools to the technological requirements. The advantage of this approach – as opposed to the usual methods – can be seen in the fact that the graphical user interface was available before the full implementation, subsequently the end users know, in advance, exactly what was being provided and how it looked. In the traditional way, the prototype develops from an idea, although the design is predetermined by the data for which the programmer has to provide the interaction. The application consists of two windows; Search Mask and Main Window. The Search Mask, by which the user can find the medical data, is displayed immediately on starting the application. Subsequently, the application switches to the main window to display the data.
The search pop-up is a self-opening search form, in which the user can enter parameters for a search in the data base.
The Viewer is a window containing the actual image data within the main window. An individual viewer window is opened for every patient or patient study.
The Player is a window, within which the selected images can be played, started/stopped, navigated etc.
The Toolbox contains all the tools (functions), which the user requires to manipulate the images.
The Symbol Toolbar is the window beneath the menu. It contains all other tools, which are not in the Toolbox, for example: grid alignment, navigation, etc.
Level 4: Implementation: From the viewpoint of Software Engineering the most essential specifications were: Platform independence (Mac, Windows, Unix, Linux); Support for the most important display formats, including DICOM, BMP, JPEG, GIF, PNG, TIFF, AVI, MPEG 4 – including this additional requirements the viewer was developed (and is still in use (November, 2013), (Holzinger, Geierhofer, Ackerl & Searle, 2005b).



Slide 12-27: Hi-Fi Prototype allows low-level interaction
It is important to test the functionality with a hi-fi prototype, either by a part implementation or with a simple assistive technology (even e.g. Power Point slides).

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)

A. Holzinger 709.049 43/88 Med Informatics L12

Your applications are used by end users – your customers – consequently a solid validation, verification and evaluation, and/or experimental examination is invaluable (for details please refer to (Holzinger, 2010)).
Let us just clarify some definitions at first:
Validation is the process of checking if and to what extend your system meets the specifications and therefore fulfils (American English: fulfills) its intended purpose.
Verification is a quality control process that is used to evaluate whether and to what extend your system complies with official regulations, legal specifications, standards or norms.
Evaluation is the systematic assessment of your application by use of certain criteria against a defined set of standards.
Experimental Examination is testing the system against stated Hypotheses (e.g. “By use of the system A the task X is performed in shorter time than by use of system B”) either in a laboratory or, better, in the field (real life experiment, field experiment).
Evaluation is the systematic assessment of your application by use of certain criteria against a defined set of standards.
Experimental Examination is testing the system against stated Hypotheses (e.g. “By use of the system A the task X is performed in shorter time than by use of system B”) either in a laboratory or, better, in the field (real life experiment, field experiment).

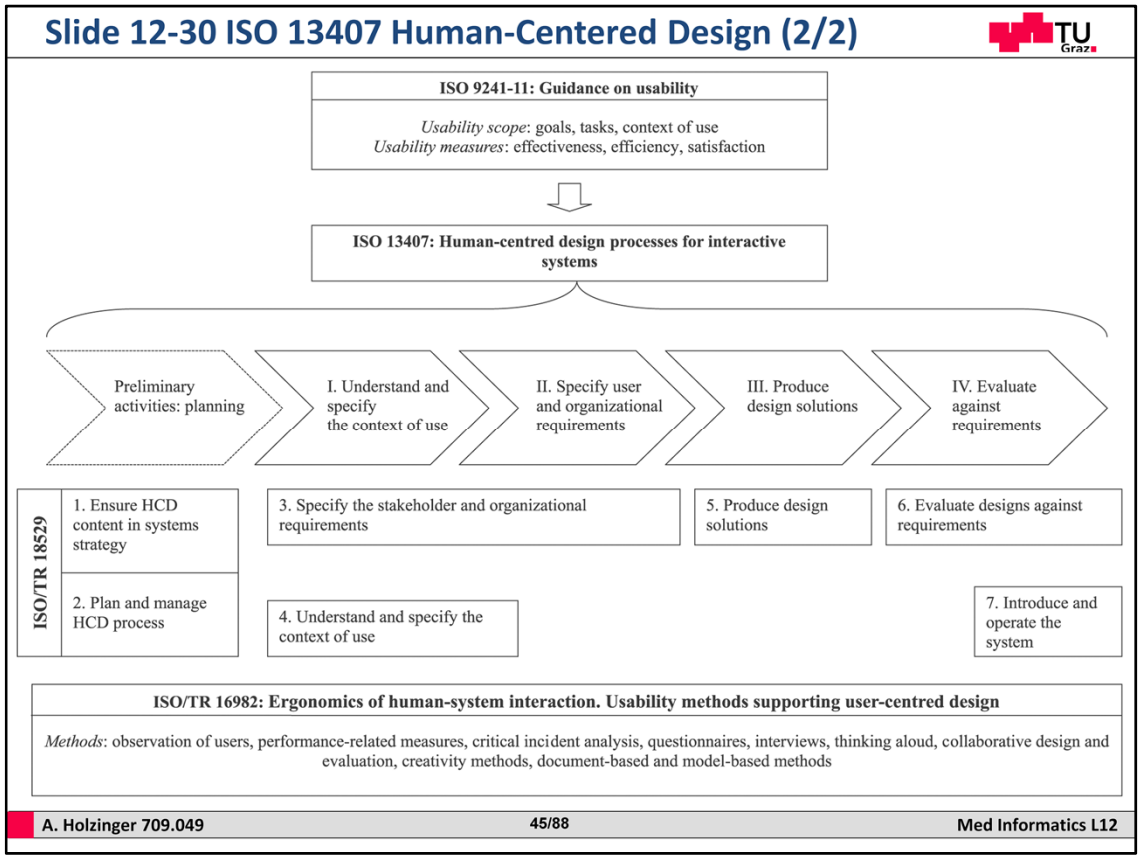
Slide 12-29 ISO 13407 Human-Centred Design (1/2)

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.*

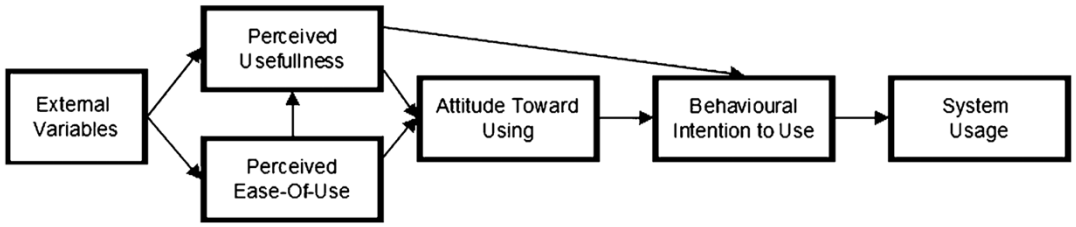
A. Holzinger 709.049 44/88 Med Informatics L12

Slide 12-29 ISO 13407 Human-Centred Design (1/2)
This brings us back to the ISO 13407 Human-Centered Design: along with ISO TR 18529 these standards represent a maturing of the discipline of user-centred design. The systems development community sees that Human Factors has processes which can be managed and integrated with existing project processes. This internationally accepted set of processes provides a definition of the capability that an organization must possess in order to implement such user-centred design effectively. It can also be used to assess the extent to which a particular development project employs user-centred design (Earthy, Jones & Bevan, 2001).



In this slide we see the “big picture” – a good overview on the contents of the ISO 13407 standard.

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.

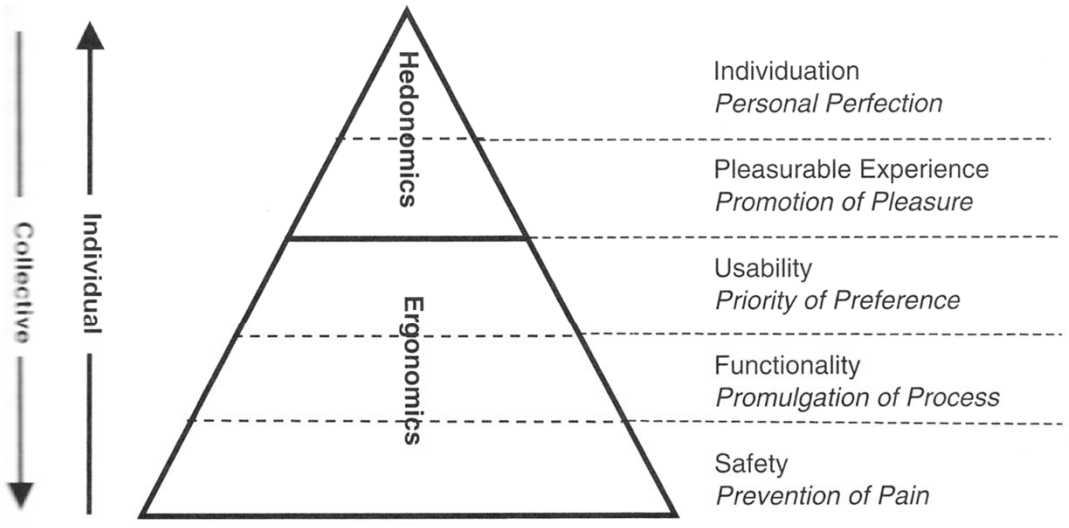
*) Davis, F. D. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 13, (3), 319-339.

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

A. Holzinger 709.049 46/88 Med Informatics L12

Originally, the framework of (Shackel, 1991) has been one of the most influential paradigms for conceptualizing the acceptability of any given system to its intended end users: He suggested that systems acceptability can be defined as a function of three orthogonal dimensions, which he balanced against cost (Holzinger, Searle & Wernbacher, 2011):
1) utility (whether the system does what is needed functionally);
2) usability (whether and to what extent the users can actually work with the system successfully); and
3) likeability (whether the users feel the system is suitable).
In the slide we see a previous model, proposed to explain and predict user acceptance: the technology acceptance model (TAM) by Davis (1989), confer also to (Davis, 1993) and (Morris & Turner, 2001). There have been several theoretical models developed in order to study user acceptance; many of them incorporate perceived ease of use as a determinant of acceptance, the Technological Acceptance Model, TAM as can be seen in this slide is most widely accepted.
Background: Originally, TAM was adapted from the Theory of Reasoned Action (TRA) by (Fishbein & Ajzen, 1975) and it proposes that two specific beliefs:
a) the perceived ease of use and
b) the perceived usefulness are determining a person's behavioral intention to use technology.
However, the attitude towards using a technology was originally omitted in the final model, due to a partial mediation of the impact of beliefs on intention by attitude, a weak direct link between perceived usefulness and attitude, and a strong direct link between perceived usefulness and intention; this was explained as originating from people intending to use technology, due to it was useful for them even though they did not have a positive affect (attitude) towards using. This derives from the Hedonomics-Ergonomics pyramide (see next slide):

Slide 12-32 Ergonomics versus Hedonomics

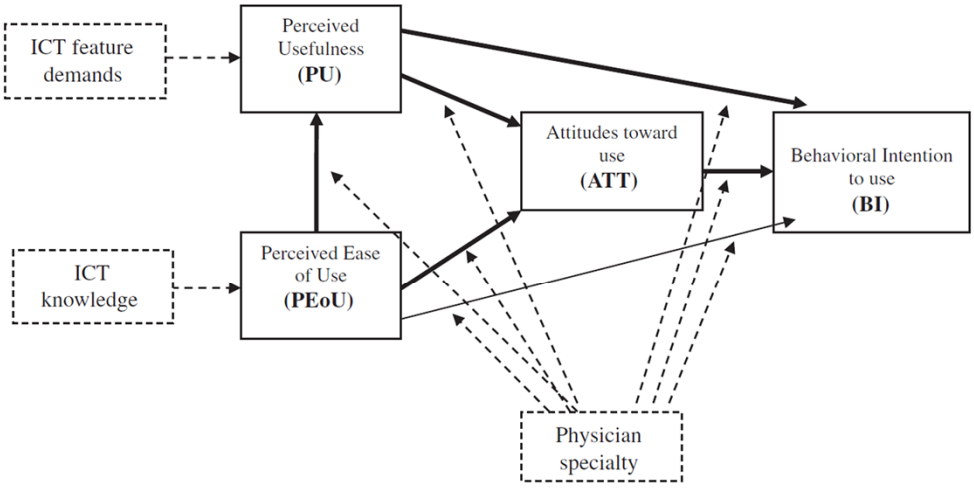


Helander, M. G. & Khalid, H. M. (2006) Affective and Pleasurable Design. In: Salvendy, G. (Ed.) *Handbook of Human Factors and Ergonomics, Third Edition. Hoboken (NJ), Wiley.*

A. Holzinger 709.049 47/88 Med Informatics L12

Slide 12-32 Ergonomics versus Hedonomics
Similar to the famous “Maslow Pyramide” (Maslow, Frager & Fadiman, 1970), Helander & Khalid (2006) proposed a Ergnomics/Hedonomics Pyramide which can be seen in this slide: The more to the top, the more individuation, in a sense of personal perfection, takes place. Safety is the basis, followed by functionality, and usability. On these ergonomic “must-haves”, there are hedonomics factors including pleasurable experiences.

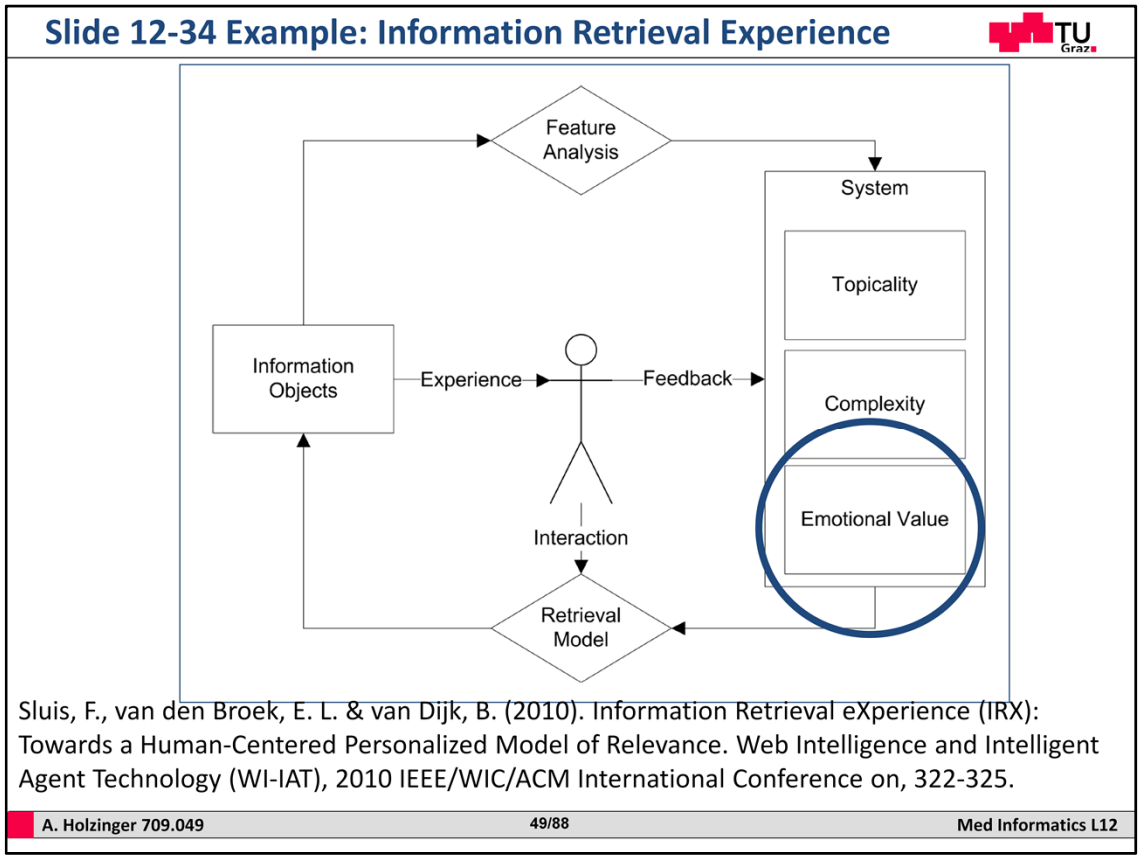
Slide 12-33 Technology Acceptance in the clinical context



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.*

A. Holzinger 709.049 48/88 Med Informatics L12

Slide 12-33 Technology Acceptance in the clinical context
In this slide we see a version of the TAM model adapted to the clinical context: Recent empirical research has utilized the TAM to advance the understanding of medical doctors' and nurses' technology acceptance in the clinical workplace. However, the majority of the reported studies are either qualitative in nature or use small samples of medical staff. Additionally, in very few studies moderators are either used or assessed despite their importance in TAM based research. The study by (Melas et al., 2011) focused on the application of TAM in order to explain the intention to use clinical information systems, in a random sample of 604 medical staff (534 physicians) working in 14 hospitals in Greece. The authors introduce physicians' specialty as a moderator in TAM and test medical staff information and communication technology (ICT) knowledge and ICE feature demands, as external variables. The results showed that TAM predicts a substantial proportion of the intention to use clinical information systems (Melas, Zampetakis, Dimopoulou & Moustakis, 2011)
A main contribution here is that it is accepted that there are diverse clinical specialists – every of them having different needs, goals and requirements. Note that this work is from 2011 sometimes easy things take very long.



Slide 12-34 Example: Information Retrieval Experience

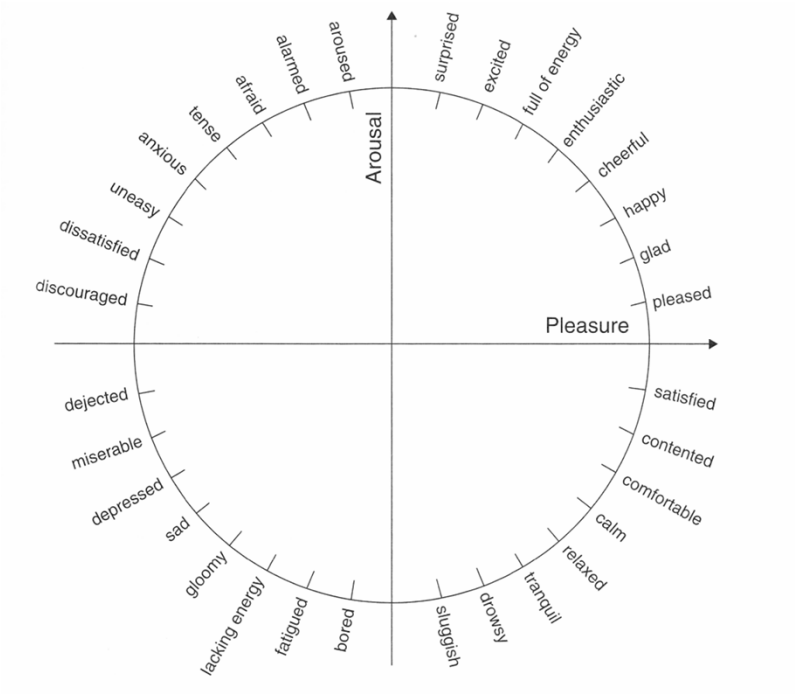
Along with technological advances, quality of use will become more important in the future including (van der Sluis, van den Broek & van Dijk, 2010):

1) Aesthetic and hedonic factors (e.g., beauty, enjoyment, and extending one's personal knowledge and satisfaction);

2) Emotional factors, addressing the antecedents and consequences of, ideally, positive emotions. Although overlapping with the first category, these factors are not seen as a goal on their own; however, they can aid in solving an information need.

3) Experiential factors, combining all contextual and related factors, including e.g., mood, expectations, and active goals) interact with the situation and time in creating the experience.

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.*

A. Holzinger 709.049 50/88 Med Informatics L12

Emotion as a trade-off between Arousal and Pleasure, and Dis-Arousal and Dis-Pleasure of course

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.

A. Holzinger 709.049 51/88 Med Informatics L12

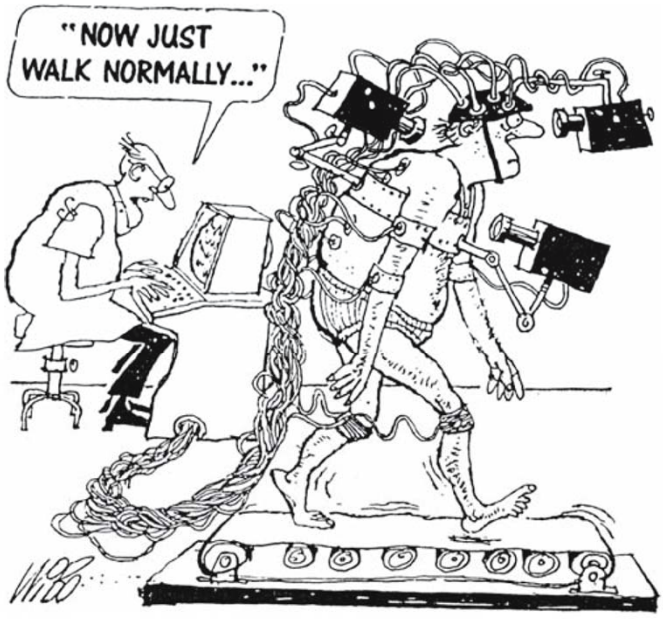
Measuring emotions is not easy and there are three basic approaches (Lopatovska & Arapakis, 2011):
1) Neuro-physiological, e.g. brain activity, pulse rate, blood pressure, skin conductance, etc. Can detect short-term changes not measurable by other means; Reliance on non-transparent, invasive sensors; 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.
2) Observation, e.g. facial expressions; speech; gestures Use of unobtrusive techniques for measuring emotion; crosscultural universals Cannot perform contextdependent interpretation of sensory data; highly dependent on environmental conditions (illumination, noise, etc.); some responses can be faked; recognises the presence of emotional expressions, not necessarily emotions.
3) 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.

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 12-37 Example methods for measuring emotion
Here just a selection of possible methods:
1) 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));
2) 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);
3) Psychogalvanic measures -> Galvanic skin response (e.g. Larson & Fredrickson (1999), Wearable sensors (e.g. Picard (2000);
4) 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

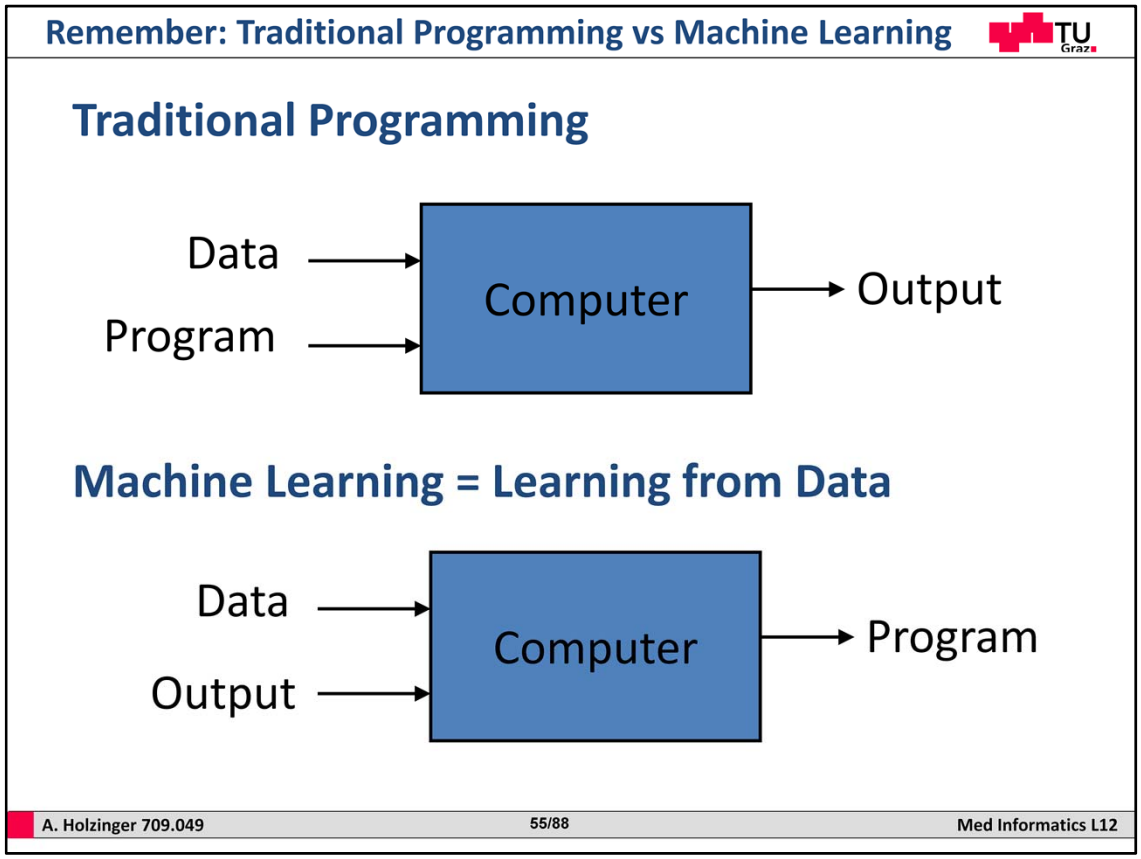


Ouwerkerk, M., Pasveer, F. & Langereis, G. (2008) Unobtrusive Sensing of Psychophysiological Parameters: Some Examples of Non-Invasive Sensing Technologies. In: Westerink, J. H. D. M. (Ed.) *Probing Experience. Heidelberg, Berlin, New York, Springer, 163-193.*

The main problem is to measure unobtrusively

Evaluation

Ok, now lets focus on evaluation issues
Evaluation is a **systematic** assessment of merit, worth, significance etc. using **criteria** benchmarked against **standards**.



As you know hard-coding problems is a bottleneck, now the idea is: Let the data do the work instead!
Automating automation: Getting computers to program themselves

Better data is often more useful than simply more data (quality over quantity)

Data collection may be expensive
Cost of time and materials for an experiment
Cheap vs. expensive data
Raw images vs. annotated images

Want to collect best data at minimal cost
http://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/

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.

A. Holzinger 709.049 56/88 Med Informatics L12

ML algorithms should be fast both in to train and application, accurate, scalable, interpretable and as simple as possible;
This sounds like an eierlegende Wollmilchsau jack of all trade – all-in-one-suitable-for-all-purposes – that this is not simple should simply be clear!

Occam's razor has been interpreted in two different ways:
1) simplicity is a goal in itself is essentially correct,
2) Simplicity leads to greater accuracy is problematic.

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.

NFL states that for certain types of mathematical problems, the computational cost of finding a solution, averaged over all problems in the class, is the same for any method. **No solution offers a 'short cut'.**

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.

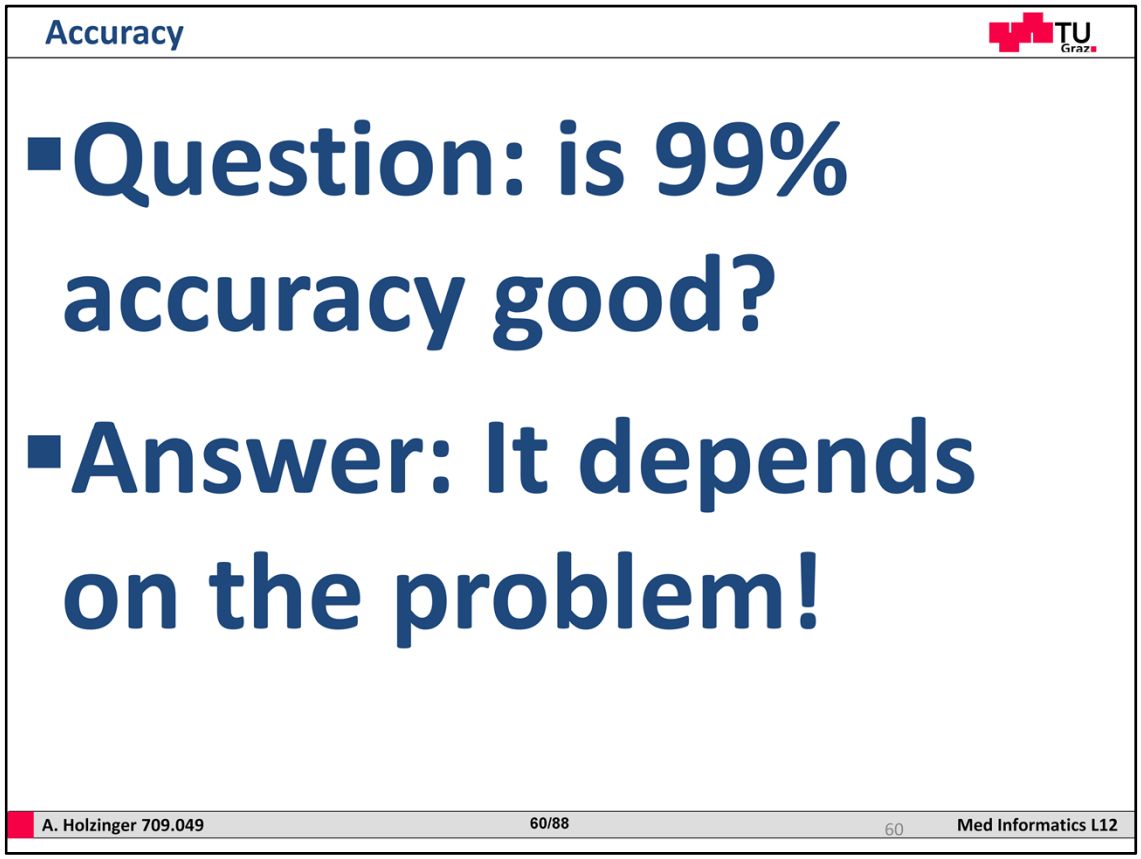
Source: Turban et al. (2011), Decision Support and Business Intelligence Systems 58

Receiver operating characteristics (ROC) graphs are useful for organizing classifiers and visualizing their performance. ROC graphs are commonly used in medical decision making, and in recent years have been used increasingly in machine learning and data mining research. Although ROC graphs are apparently simple, there are some common misconceptions and pitfalls when using them in practice.
The purpose of this article is to serve as an introduction to ROC graphs and as a guide for using them in research

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/



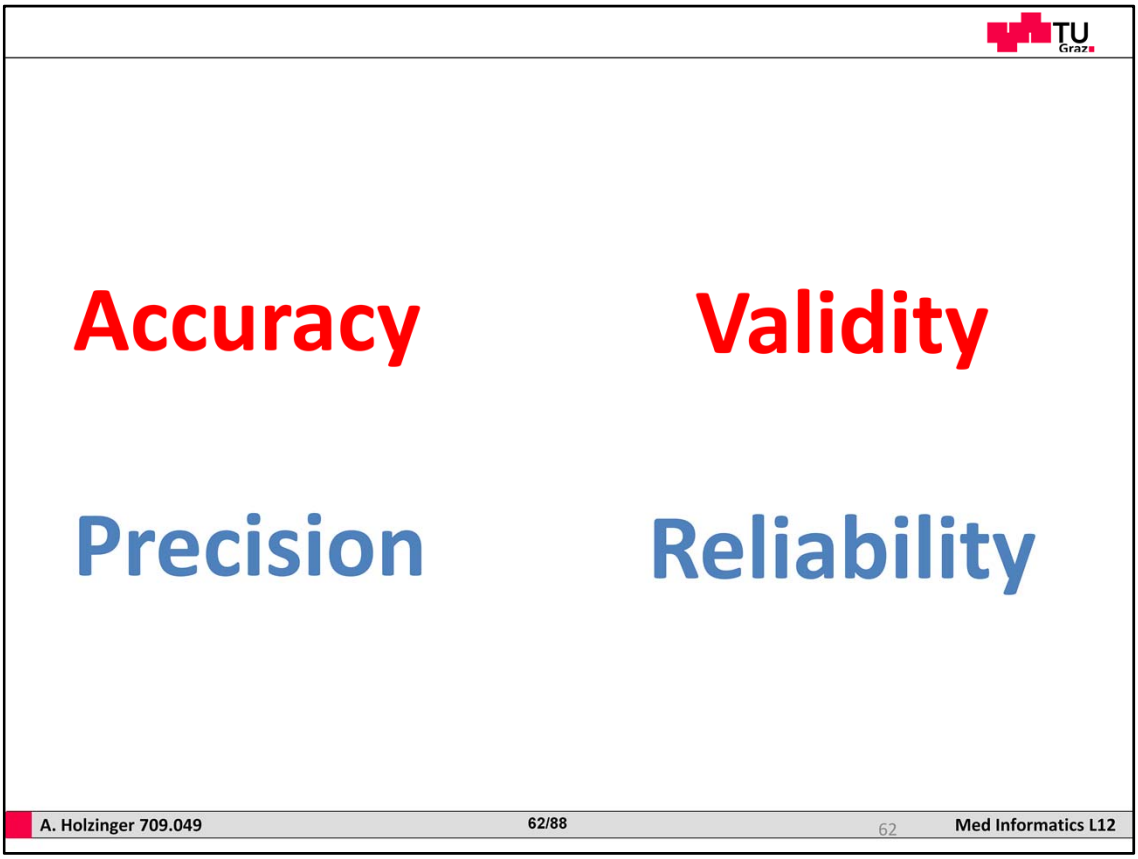
99% accuracy good?
–can be excellent, good, mediocre, poor, terrible
–depends on problem
• is 10% accuracy bad?
–information retrieval
• BaseRate = accuracy of predicting predominant class
(on most problems obtaining BaseRate accuracy is easy)

We focus now on four measures:

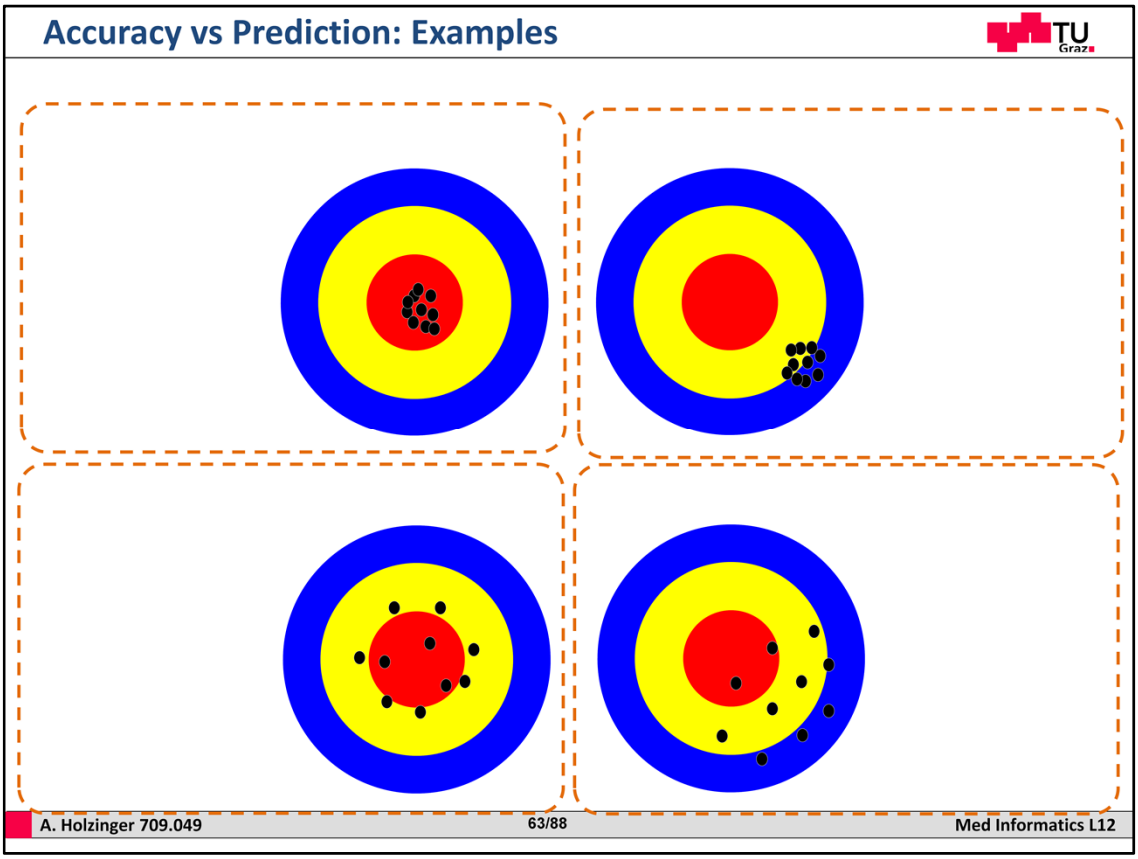
Please always remember these four terms:

- **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

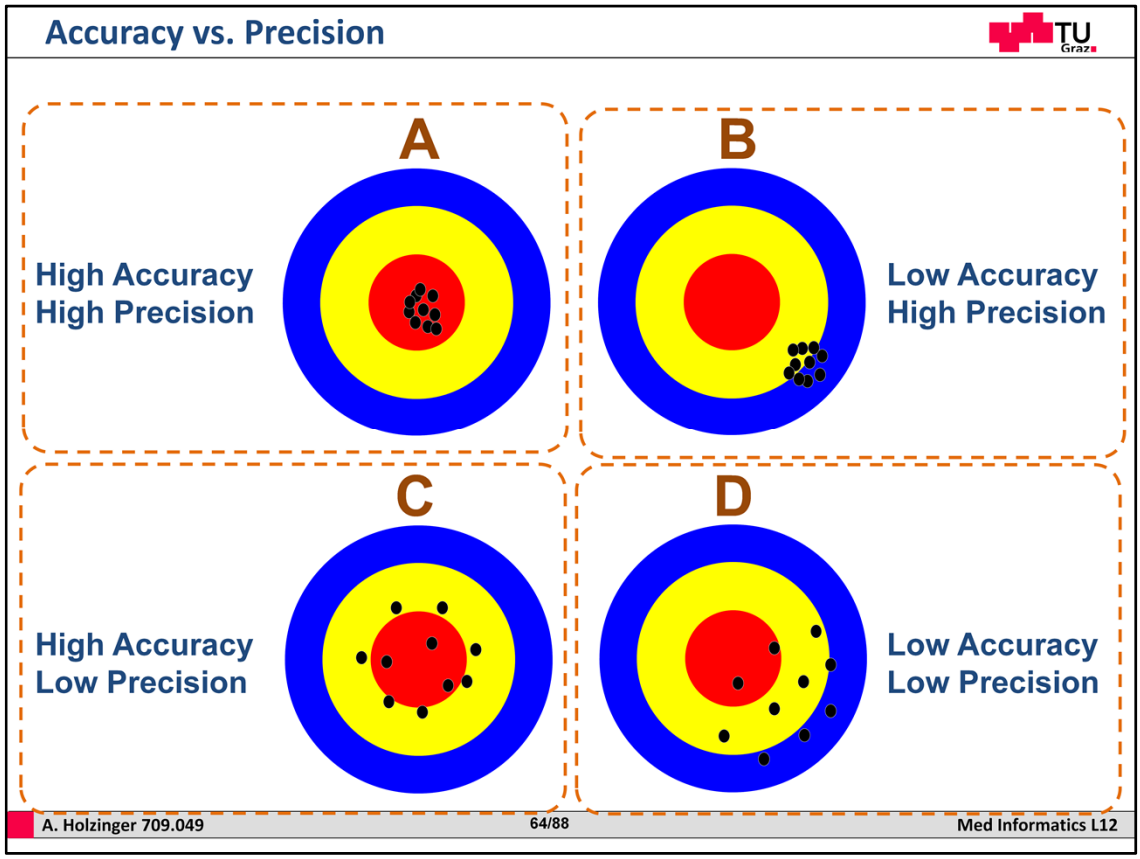
Accuracy is usually estimated by using an independent test set that was not used at any time during the learning process. More complex accuracy estimation techniques, such as cross-validation and the bootstrap, are commonly used, especially with data sets containing a small number of instances.


Accuracy
Validity
Precision
Reliability
A. Holzinger 709.049
62/88
62
Med Informatics L12

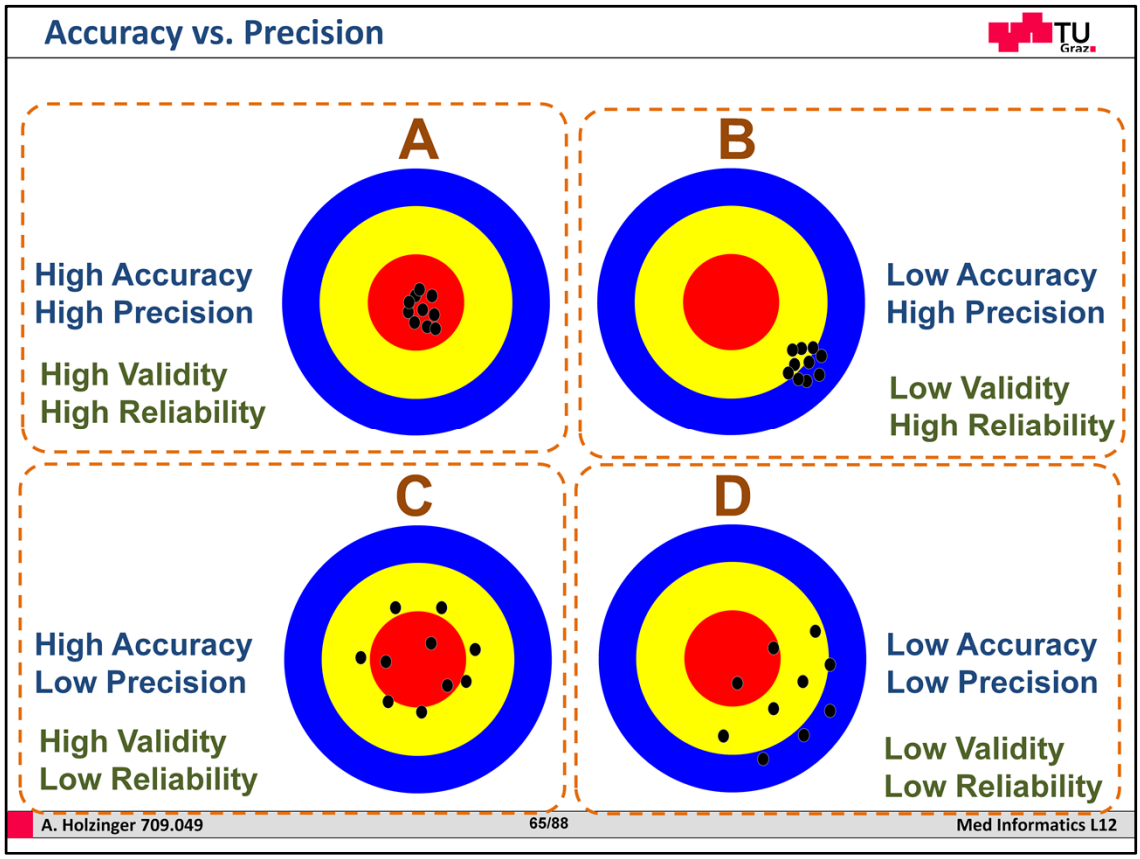
Accuracy

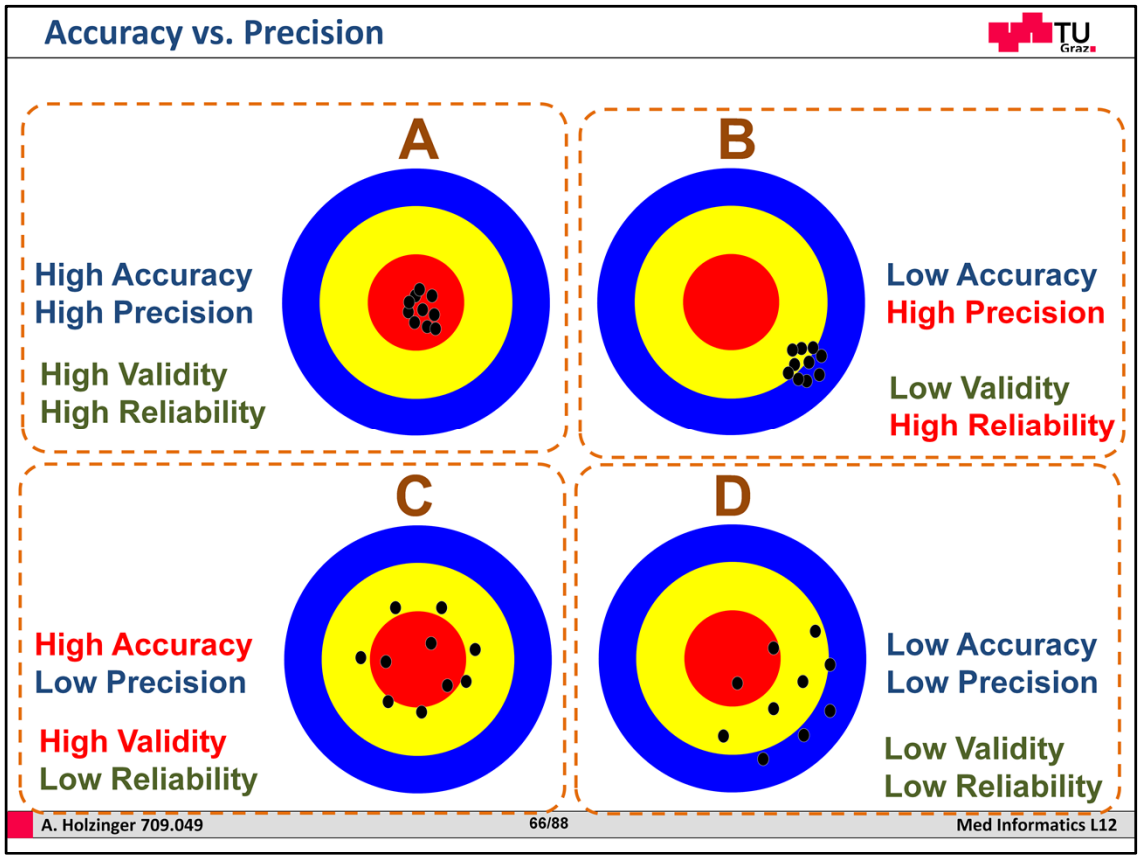


Let us assume we have four cases: A, B, C, D



Case A: high accuracy, high precision – super!
Case B: not quite so super
Case C: not so good
Case D: worst case


Accuracy vs. Precision
TU Graz
A
High Accuracy
High Precision
High Validity
High Reliability
B
Low Accuracy
High Precision
Low Validity
High Reliability
C
High Accuracy
Low Precision
High Validity
Low Reliability
D
Low Accuracy
Low Precision
Low Validity
Low Reliability
A. Holzinger 709.049
65/88
Med Informatics L12


Accuracy vs. Precision
TU Graz
A
High Accuracy
High Precision
High Validity
High Reliability
B
Low Accuracy
High Precision
Low Validity
High Reliability
C
High Accuracy
Low Precision
High Validity
Low Reliability
D
Low Accuracy
Low Precision
Low Validity
Low Reliability
A. Holzinger 709.049
66/88
Med Informatics L12

Please remember:

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

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

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

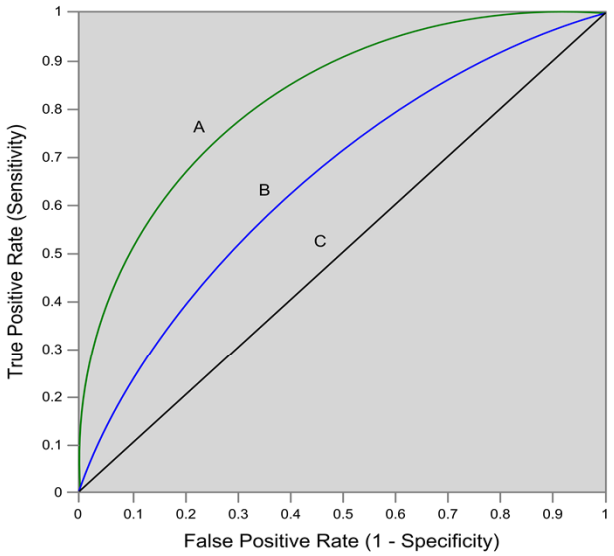
$$True\ Negative\ Rate = \frac{TN}{TN + FP}$$

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

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

True Positive Rate is called sensitivity!
True Negative Rate is called specifcity! Pronounce:

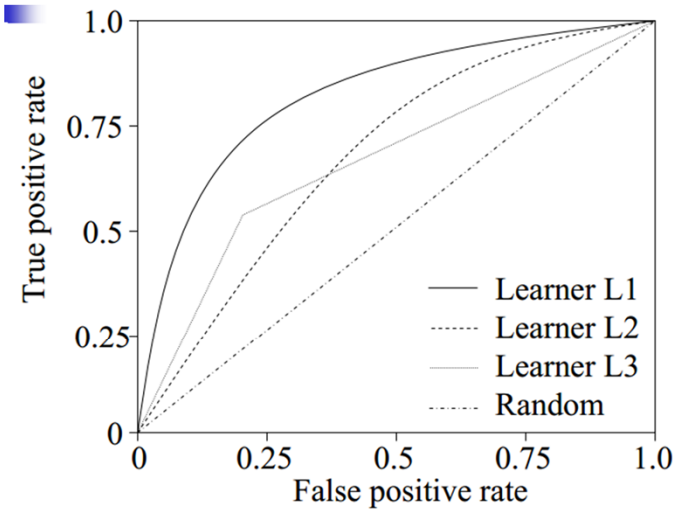
Again the ROC Curve



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.

Source: Turban et al. (2011), Decision Support and Business Intelligence Systems 68

Examples



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

Learner L1 dominates L2 is L2's ROC curve is beneath L1's curve
If L1 dominates L2, then L1 better than L2 for all possible costs and class distributions
If neither dominates (L2 and L3), then there are times when L2 maximizes accuracy, but does not minimize cos

Future Outlook

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

Source: Ham & Kamber (2000) 70

A last word …

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

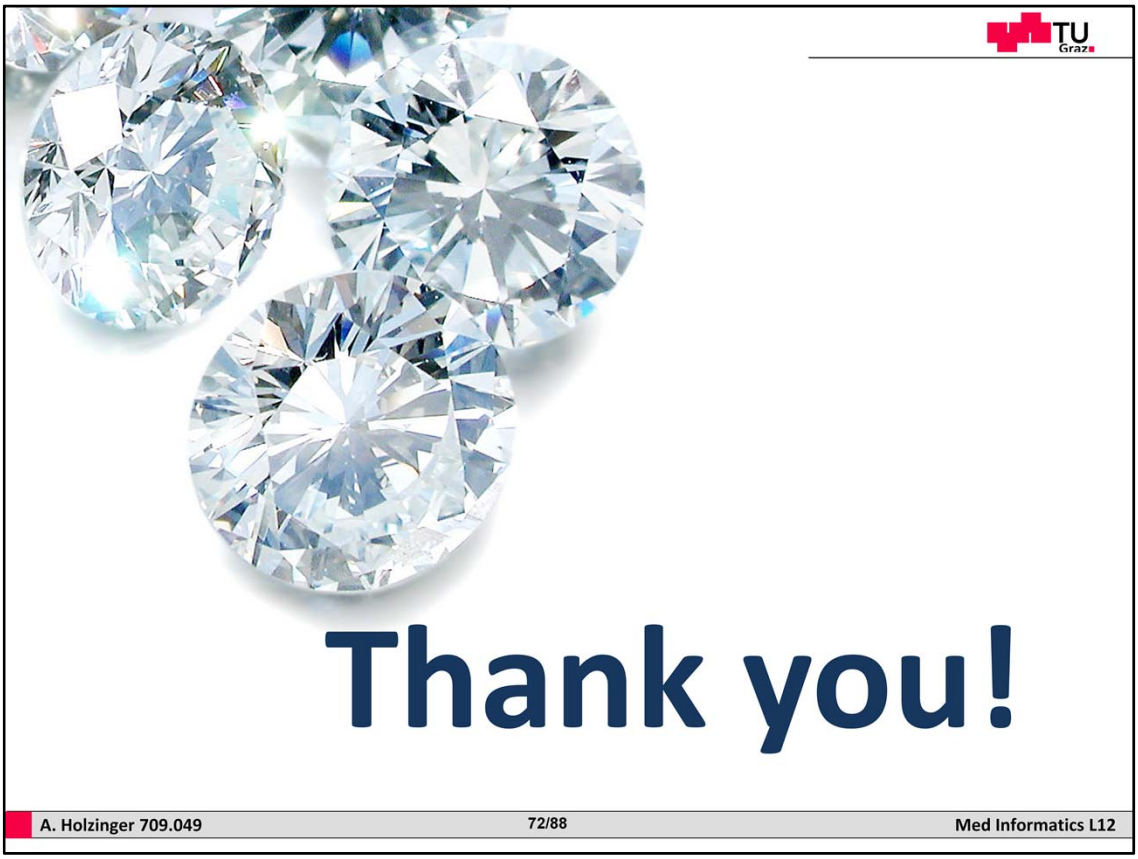
Lopez-Paz, D., Muandet, K., Schölkopf, B. & Tolstikhin, I. 2015. Towards a learning theory of cause-effect inference. Proceedings of the 32nd International Conference on Machine Learning, JMLR, Lille, France.

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

A. Holzinger 709.049 71/88 Med Informatics L12

In the Tudor Period in the same year Elizabeth I was born – Oil on oak painting. Given a visual puzzle – there can be seen a weird object in the front – and only from the side the skull is visible.

This is a good example for the discovery of causal relationships from purely observational data, which is a fundamental problem in science.



My DEDICATION is to make data valuable … 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!

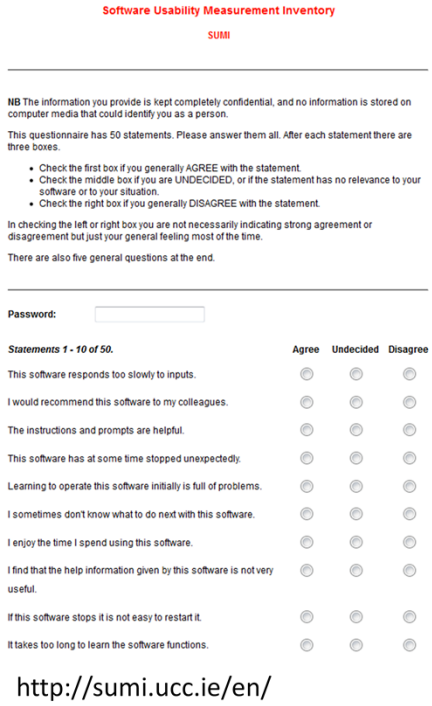
Sample Questions (2/2)

- 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



Software Usability Measurement Inventory

SUMI

NB The information you provide is kept completely confidential, and no information is stored on computer media that could identify you as a person.

This questionnaire has 50 statements. Please answer them all. After each statement there are three boxes.

- Check the first box if you generally AGREE with the statement.
- Check the middle box if you are UNDECIDED, or if the statement has no relevance to your software or to your situation.
- Check the right box if you generally DISAGREE with the statement.

In checking the left or right box you are not necessarily indicating strong agreement or disagreement but just your general feeling most of the time.

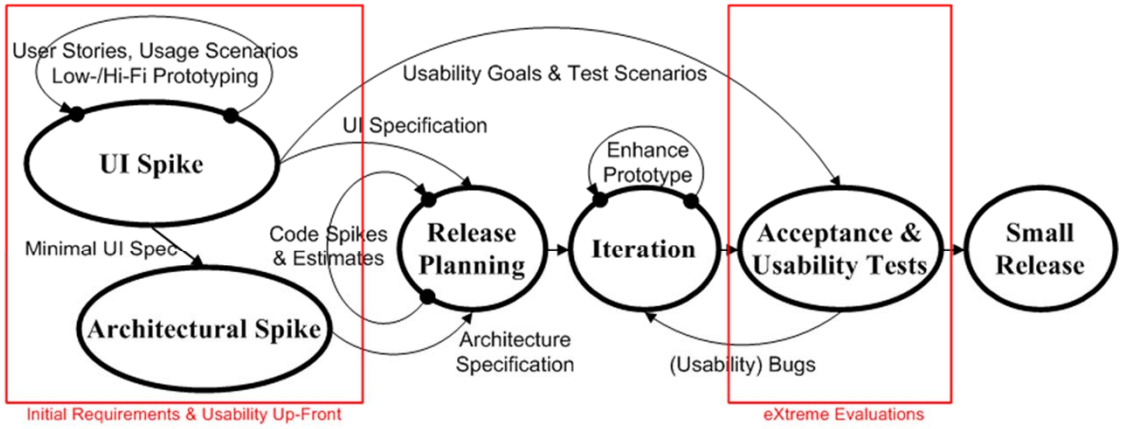
There are also five general questions at the end.

Password:

Statements 1 - 10 of 50.	**Agree**	**Undecided**	**Disagree**
This software responds too slowly to inputs.			
I would recommend this software to my colleagues.			
The instructions and prompts are helpful.			
This software has at some time stopped unexpectedly.			
Learning to operate this software initially is full of problems.			
I sometimes don't know what to do next with this software.			
I enjoy the time I spend using this software.			
I find that the help information given by this software is not very useful.			
If this software stops it is not easy to restart it.			
It takes too long to learn the software functions.			

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.*

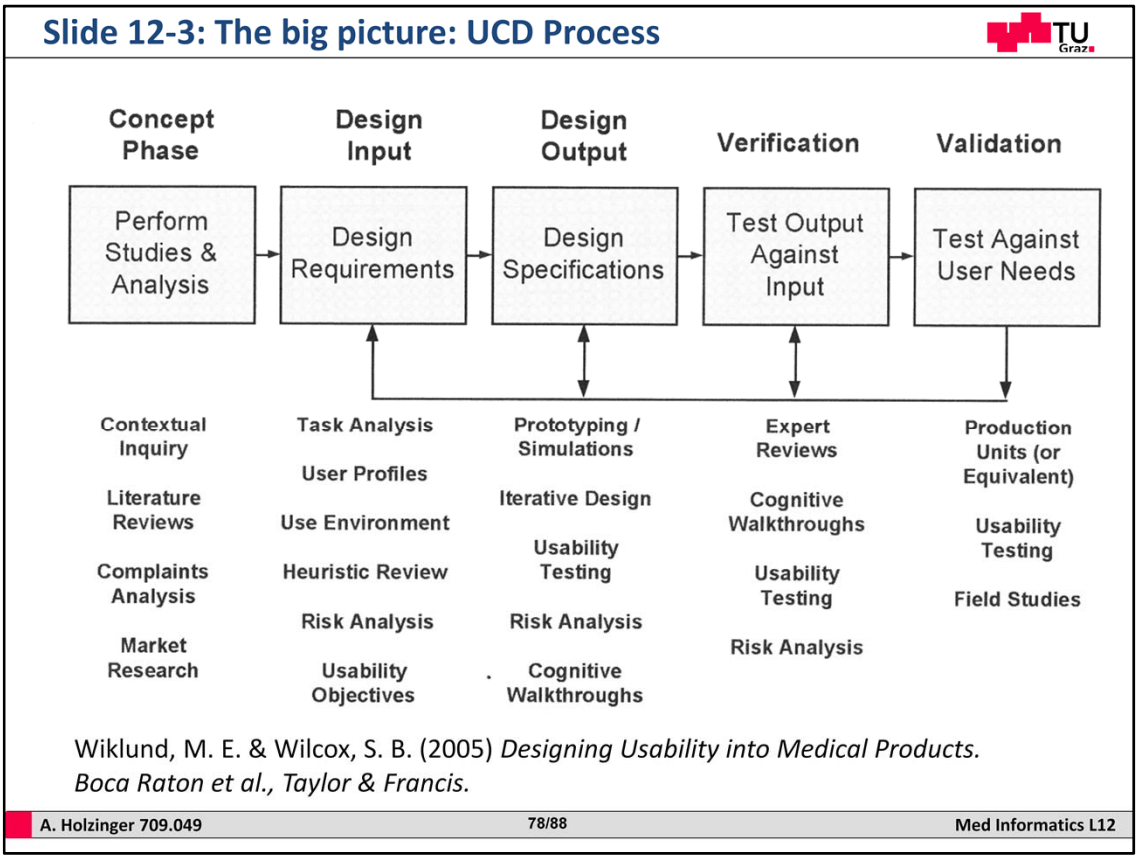
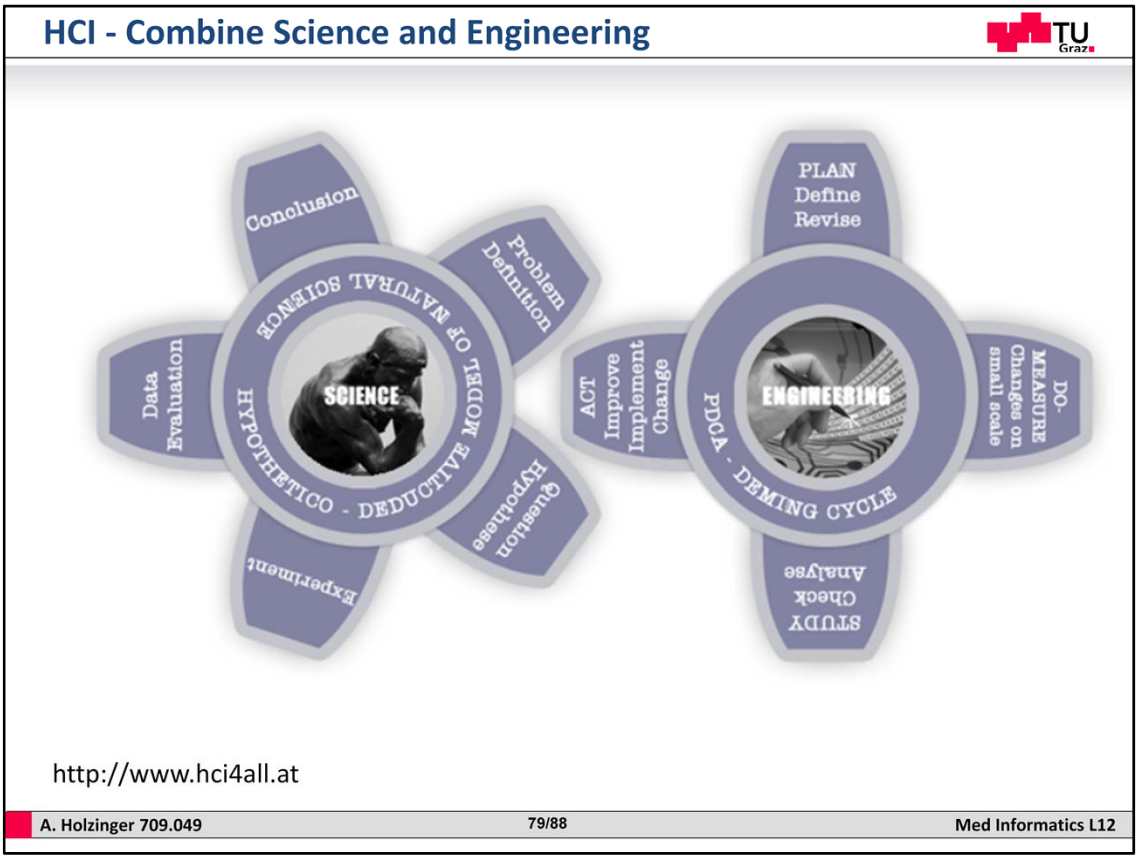

Slide 12-3: The big picture: UCD Process
TU Graz
Concept Phase
Design Input
Design Output
Verification
Validation
Perform Studies & Analysis
Design Requirements
Design Specifications
Test Output Against Input
Test Against User Needs
Contextual Inquiry
Literature Reviews
Complaints Analysis
Market Research
Task Analysis
User Profiles
Use Environment
Heuristic Review
Risk Analysis
Usability Objectives
Prototyping / Simulations
Iterative Design
Usability Testing
Risk Analysis
Cognitive Walkthroughs
Expert Reviews
Cognitive Walkthroughs
Usability Testing
Risk Analysis
Production Units (or Equivalent)
Usability Testing
Field Studies
Wiklund, M. E. & Wilcox, S. B. (2005) Designing Usability into Medical Products. Boca Raton et al., Taylor & Francis.
A. Holzinger 709.049
78/88
Med Informatics L12

Fig 26.1., p.260



Quality Approach: PDCA Deming Wheel
This concept was developed by (Shewhart, 1958) as PDSA cycle. The roots can be tracked back to Aristotle (384–322 BC) and Francis Bacon (1561–1626). The PDSA cycle consists of four steps:
1) PLAN: Study the process;
2) DO: Make changes on a small scale;
3) STUDY: Observe the effects and
4) ACT: Identify what you can learn from your observation.
William E. Deming (1900–1993) promoted this model effectively and called it PDCA cycle (Deming, 1994) and is also known as Deming wheel:
1) PLAN: Clearly define the objectives and processes necessary to gain deliverables in accordance with the expected output;
2) DO: Implement the new processes on a small scale (e.g. within a trial or pilot project);
3) CHECK: Now measure the outcome and compare your results against the expected results and look for differences;
4) ACT: Finally, analyze the differences to determine their cause.
Each finding can be used as input for a new PDCA cycle.
The PDCA wheel can be used to coordinate your continuous improvement. Every improvement starts with a goal and with a plan on how to achieve that goal, followed by action, measurement and comparison of the gained output. The most important issue is that you act – on a small scale – but act.
Remember the “Write now!” approach in section 4.4.
Deming introduced a “System of Profound Knowledge”, consisting of four parts (Stepanovich, 2004):

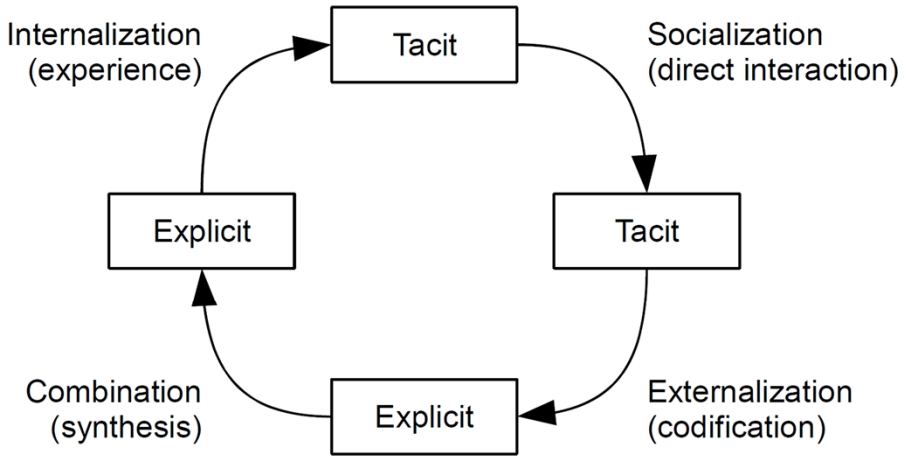
Comparison of Usability Engineering Methods

	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	Factors									
	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	+			+					+	+

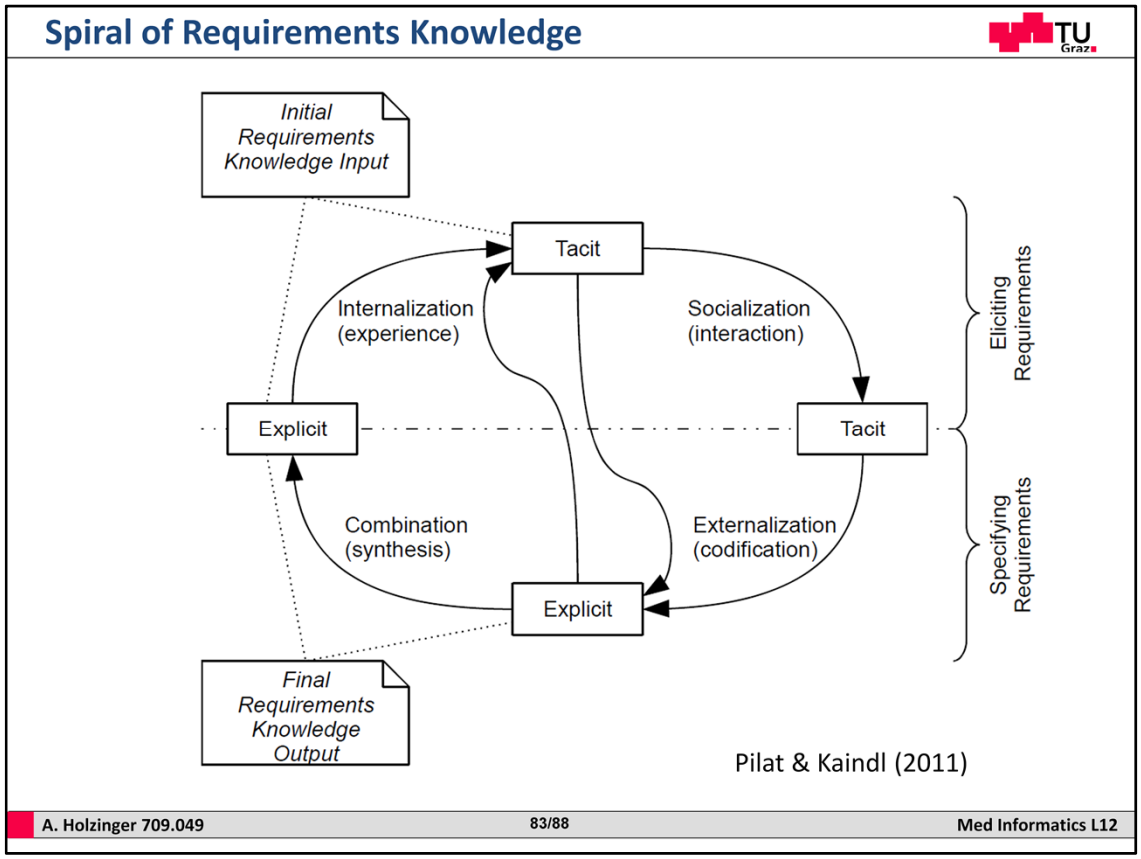
Remember: Cyclic View of Nonaka's Spiral of Knowledge



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

A. Holzinger 709.049 82/88 Med Informatics L12

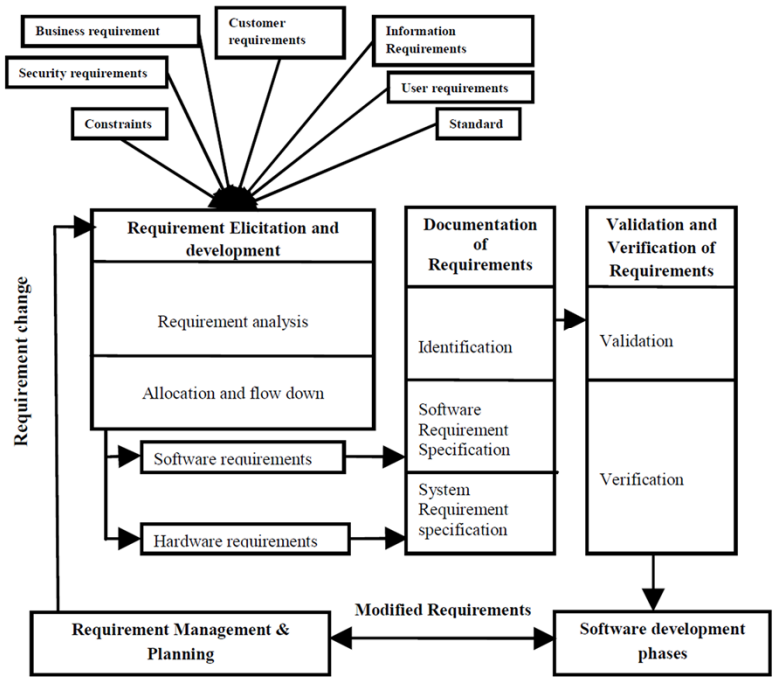
A good example for the application of Nonaka & Takeuchi Knowledge Spiral.



For RE, it has been recognized that the knowledge of stakeholders is mostly tacit. Therefore, the initial knowledge
available when doing requirements will also be primarily tacit.

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.*

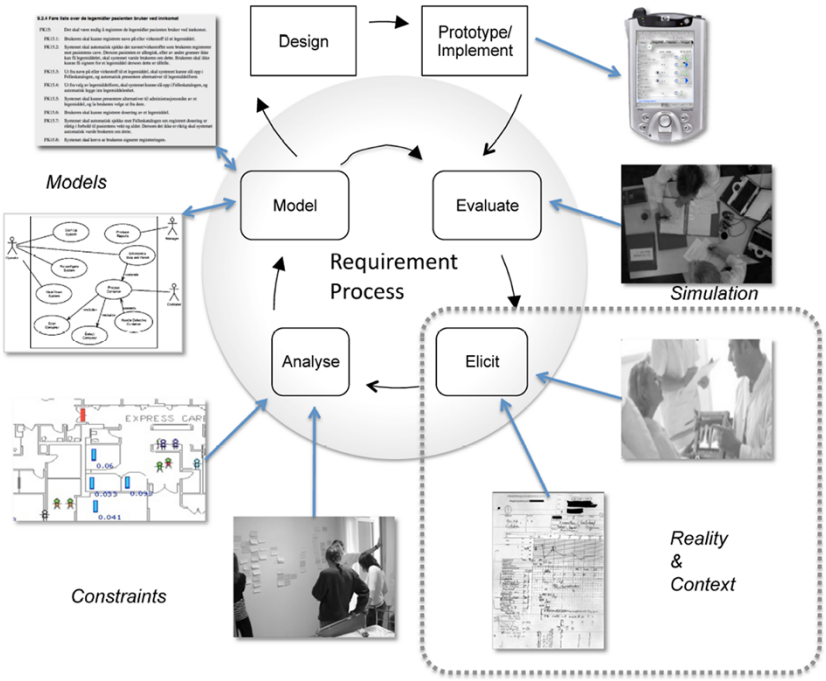


A. Holzinger 709.049 84/88 Med Informatics L12

Just an Example

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.*



A. Holzinger 709.049 85/88 Med Informatics L12

A good example for requirements engineering.

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)

A. Holzinger 709.049 86/88 Med Informatics L12

In 2007, the authors did a comprehensive study of clinical work practice based on an earlier described approach to structured observation. Twenty-five physicians and nurses at two large Norwegian University Hospitals were followed for a total of 55 days. The purpose of the study was to investigate information and communication patterns in typical ward situations.

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.*

(19) **United States**
(12) **Patent Application Publication**
SHIN et al.
(10) **Pub. No.: US 2011/0137137 A1**
(43) **Pub. Date: Jun. 9, 2011**

(54) **SENSING DEVICE OF EMOTION SIGNAL AND METHOD THEREOF**

(75) Inventors: **Hyun-Soon SHIN**, Daejeon (KR); **Yong-Kwi Lee**, Seoul (KR); **Jun Jo**, Daejeon (KR); **Ji-Hoon Kim**, Daejeon (KR); **Jun-Sik Choi**, Daejeon (KR); **In-Tark Han**, Daejeon (KR)

(73) Assignee: **Electronics and Telecommunications Research Institute**, Daejeon (KR)

(21) Appl. No.: **12/959,214**

(22) Filed: **Dec. 2, 2010**

(30) **Foreign Application Priority Data**

Dec. 8, 2009 (KR) 10-2009-0121185

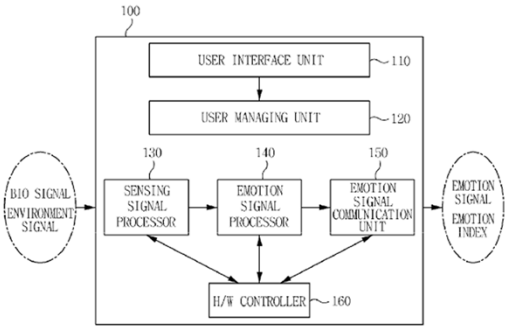
Publication Classification

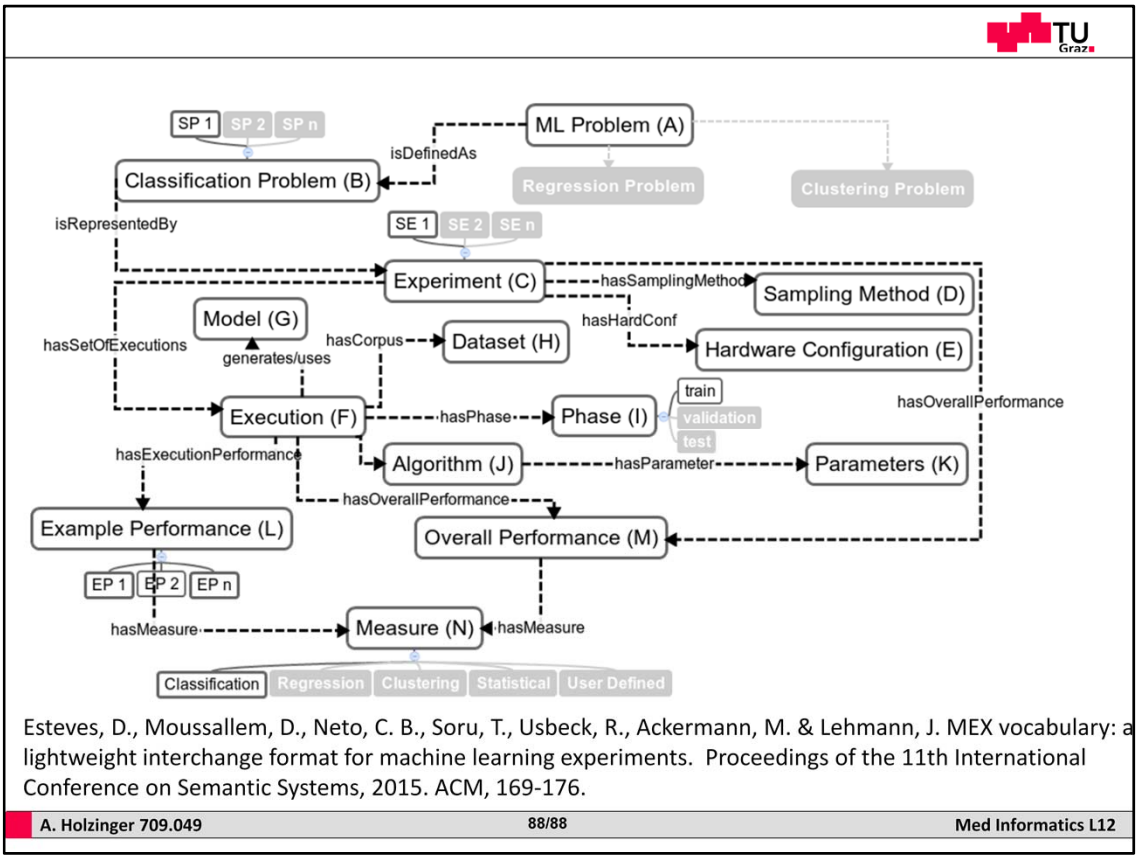
(51) **Int. Cl.**
G09B 19/00 (2006.01)
A61B 5/00 (2006.01)

(52) **U.S. Cl.** **600/301**; 434/236

(57) **ABSTRACT**

The present invention relates to a sensing device of an emotion signal and a method thereof capable of recognizing and analyzing the change in emotion by collecting at least one of the bio signals and peripheral environmental signals, guiding the state of emotion to a user, and sharing emotion information with authorized persons and communicating it between them. The sensing device of the emotion signal according to the embodiment of the present invention includes a sensing signal processor that senses a plurality of bio signals and peripheral environment signals of a user; an emotion signal processor that generates emotion signals representing the emotional state of the user for each of the sensed bio signals and collects the emotion signals to generate an emotion index; and an emotion signal communication unit that optionally transmits the emotion signals and the emotion index to external authorized devices.





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.

The diagram representing the classical iterations for an execution of a machine learning problem (Classification, Regression or
Clustering). The white rounded rectangles representing a complete path for a Classification problem as well as its input (Model, Corpus,
Phase and Algorithm) and outputs variables (Example Performance and Overall Performance)