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

- understand the concepts and importance of usability
- be 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
- be familiar with 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
- be 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), and again the ROC

Keywords of the 12th Lecture

- Action analysis/Cognitive walkthrough
- Emotion recognition
- Ergonomics
- Human factors
- Human-centered designs (HCD)
- Medical Device Directive (MDD)
- Medical Product Law
- Medical Software
- Medizin-Poduzte-Gesetz (MPG)
- Usability
- Technology Acceptance Model (TAM)
- Thinking aloud
- Usability Engineering (UE)
- User-Centred Design (UCD)
- Verification

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
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]

References:


Slide 12.2: Medical Workplace Usability - enhance quality


Please remember:

Anonymization
- Secure
- Confidential
- Accessible

Pseudonymization
- Secure
- Confidential
- Accessible

Remember: Information Quality as the hiatus theoreticus
Slide 12-3: A framework for understanding usability

Slide 12-4: System characteristic versus Quality factor

Slide 12-5: ISO Standards for Healthcare

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

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

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

ISO/TC 216 Health informatics

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.

WS 2015
Quality first!


Remember: In medicine we have two different worlds …

Our central hypothesis: Information bridges this gap


Slide 12-10a The origins: Kaizen

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

Slide 12d Deming Wheel

William Edwards Deming (1900-1993)

Quality Management System

Continuous Quality Improvement

---

Slide 12-11: Quality Improvement Cycle

Plan continuous improvement
Define the system
Standardize improvements
Act
Study the results
Do
Try out improvement theory
Analyze cases
Assess current situation


---

Slide 12-12 Product vs. Process Quality

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


---

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


---

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

- Functionality: accuracy, scalability, interoperability, security
- Reliability: maintainability, fault tolerance, recoverability, availability
- Efficiency: time behaviour, resource management
- Maintainability: analysability, changeability, stability
- Portability: adaptability, interchangeability, co-existence, replaceability
- Usability: understandability, learnability, operability, attractiveness


---

Slide 12-15: Remember Medical workflows...

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

Slide 12-16: Comparison of Usability Engineering Methods

<table>
<thead>
<tr>
<th>Inspection Methods</th>
<th>Test Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heuristic Evaluation</td>
<td>Cognitive Walkthrough</td>
</tr>
<tr>
<td>Applicable in Phase</td>
<td>all</td>
</tr>
<tr>
<td>Required Time</td>
<td>low</td>
</tr>
<tr>
<td>Needed Users</td>
<td>none</td>
</tr>
<tr>
<td>Required Resources</td>
<td>3+</td>
</tr>
<tr>
<td>Required Equipment</td>
<td>low</td>
</tr>
<tr>
<td>Interactive</td>
<td>no</td>
</tr>
</tbody>
</table>


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


Slide 12-18: Software Usability Measurement Inventory (SUMI)

A funny video about SUMI can be found here:
http://www.youtube.com/watch?v=SVE2yhSyIy4


Slide 12-19 Quantifying Usability Metrics in Software Quality


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

Slide 12-21: Remember the big picture: UCD Process
The power of iteration: A UCD spiral


Agility: Make the UCD spirals as small as possible


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.

Rapid Prototyping – Paper Mock-ups


Hi-Fi Prototype allows low-level interaction

Holzinger et al. (2005)
Validation = a (external) quality process to demonstrate (to the stakeholder) that the system complies with the original specifications;

Verification = a (internal) quality process, used to evaluate whether and to what extent the system complies with the original specifications;

Holzinger et al. (2005)

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 12-36 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


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 & Csikszentmihalyi (1983); Affect Grid (e.g. Russell 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))

- **Psychophysiological measures** → Galvanic skin response (e.g. Larson & Fredrickson (1999), Wearable sensors (e.g. Picard (2000));

- **Performance measures** → Judgment task involving probability estimates (e.g. Ketaelev (1989); Lexical decision task (e.g. Challis & Krane (1988), Niedenthal & Setterlund (1994))
Remember: Traditional Programming vs Machine Learning

Traditional Programming

Data → Computer → Output

Program

Machine Learning = Learning from Data

Data → Computer → Program

Output

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 sine necessitate,” which, approximately translated, means Entities should not be multiplied beyond necessity


Slide 12-40 NFL-Theorem


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)


FYI: Datasets for benchmarking purposes

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

Accuracy

- Question: is 99% accuracy good?
- Answer: It depends on the problem!
Please always remember these four terms:

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

---

**Accuracy** vs **Precision**

- High Accuracy, High Precision
- High Accuracy, Low Precision
- Low Accuracy, High Precision
- Low Accuracy, Low Precision
Please remember:

\[
\text{True Positive (TP)} \quad \text{True Negative (TN)} \\
\text{False Positive (FP)} \quad \text{False Negative (FN)}
\]

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}


Future Outlook

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

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!

Some useful links (1)

- [http://www.measureusability.com/sus.php](http://www.measureusability.com/sus.php) (Measuring Usability with the System Usability Scale [SUS])
- [http://sumi.ucc.ie](http://sumi.ucc.ie) (Software Usability Measurement Inventory [SUMI])
- [http://www.gesetzesarchiv.gv.at/site/6274/default.aspx](http://www.gesetzesarchiv.gv.at/site/6274/default.aspx) (Österreichische Datenschutzkommission, Austrian data Protection Commission)
- [http://iaidq.org](http://iaidq.org) (The International Association for Information and Data Quality (IAIDQ))
- [http://www.iso.org/iso/iso_9000_selection_and_use.htm](http://www.iso.org/iso/iso_9000_selection_and_use.htm) (Selection and use of the ISO 9000 family of standards)
- [http://www.measuringusability.com/sus.php](http://www.measuringusability.com/sus.php) (Measuring Usability with the System Usability Scale [SUS])
- [http://sumi.ucc.ie](http://sumi.ucc.ie) (Software Usability Measurement Inventory [SUMI])
- [http://www.gesetzesarchiv.gv.at/site/6274/default.aspx](http://www.gesetzesarchiv.gv.at/site/6274/default.aspx) (Österreichische Datenschutzkommission, Austrian data Protection Commission)
- [http://iaidq.org](http://iaidq.org) (The International Association for Information and Data Quality (IAIDQ))
- [http://www.iso.org/iso/iso_9000_selection_and_use.htm](http://www.iso.org/iso/iso_9000_selection_and_use.htm) (Selection and use of the ISO 9000 family of standards)

Appendix: Agile Process Model


Appendix: Software Usability Measurement Inventory

http://www.measuringusability.com/sus.php

Appendix: 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 Hedonics 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?
HC1 - Combine Science and Engineering

Comparison of Usability Engineering Methods


Remember: Cyclic View of Nonaka’s Spiral of Knowledge


Example: Requirement Engineering Process Model

Elicitation in the requirements process in the health domain


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 patents. X-documents are problematic, because every X-document is detrimental for any further patent application in the area of the X-document.


Actors and Information Categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Value</th>
</tr>
</thead>
</table>
| Patient information | Biographical data (BRD)  
|  Family/social history (FSH)  
|  Residency of patient  
|  Previous history (PREH)  |
| Alos | Allergies  
|  Rashes for referral (RFR)  
|  Previous ophthalmology (PREO)  |
| Paper based | Previous condition  
|  Present condition (PSC)  
|  Present condition (PSC)  
|  Present condition (PSC)  
|  Present condition (PSC)  |
| Proc | Electronic health record (EHR)  
|  Medical record (MED)  
|  Medical record (MED)  
|  Medical record (MED)  |
| Focus | Procedure  
|  Examination (EX)  
|  Examination (EX)  
|  Examination (EX)  |

Nytro, Sorby & Karpati (2009)