



## Especificação de Requisitos de Usabilidade

Engenharia de Usabilidade

Prof.: Clarindo Isaías Pereira da Silva e Pádua

Synergia / Gestus

Departamento de Ciência da Computação - UFMG

2



## Referências

- Hix, D.; Hartson, H. R. Developing User Interfaces: ensuring usability through product & process, John Wiley and Sons, 1993.
- Good, M. et al. User Derived Impact Analysis as a Tool for Usability Engineering, Proceedings of CHI Conference on Human Factors in Computing Systems, New York: ACM, 241-246, 1986.
- Gilb, T., Design by Objectives, Unpublished Manuscript, 1981.
- <http://lap.umd.edu/QUIS> (site QUIS).

3



## Especificação de Requisitos de Usabilidade

- Introdução
- Tabela de ERU
- Atributos de Usabilidade
- Instrumento de medida
- Valor a ser medido
- Níveis de desempenho
- Diretrizes

4



## Introdução

- Define metas quantitativas de usabilidade que são usadas como referência para se avaliar a qualidade da interface do usuário.
  - As especificações devem ser estabelecidas o mais cedo possível no processo de desenvolvimento.
- As metas são níveis de desempenho em tarefas típicas dos diversos papéis de usuários.

5



## Motivação

- A Especificação de Usabilidade indica quando o projeto de desenvolvimento está convergindo em direção a uma interface com sucesso.
  - Com o estabelecimento da Especificação de Usabilidade o mais cedo possível no processo de desenvolvimento, e monitorando-as a cada iteração, pode-se determinar quando a interface está, de fato, indo em direção à melhorias.

6



- Engenharia de Usabilidade é um processo através do qual as características de usabilidade são especificadas, antecipadamente e de forma quantitativa no processo de desenvolvimento, e medidas durante todo o processo (Good et al. 1986).
- Sem especificações mensuráveis, é impossível determinar metas de usabilidade e dizer se o produto final alcança essas metas. No fundo, “...se você não pode realizar medidas em uma atividade, você provavelmente não pode gerenciá-la.”

7



## Tabela de ERU

- O conceito de especificação formal de atributos em formato de tabela foi desenvolvida por Gilb (1981).
- A tabela ERU é uma das principais fontes para o planejamento das avaliações com os usuários.
  - Deve ser utilizada tanto para avaliações formativas, aquelas realizadas durante o desenvolvimento do software, quanto somativas, ou seja, avaliações de produtos já existentes.

8



### Exemplo de tabela ERU

Ordem	Identificador	Requisito ou meta?	Atributo de usabilidade	Ator	Instrumento de medida	Valor a ser medido	Níveis de desempenho			
							Atual	Pior aceitável	Melhor Possível	Alvo
1	RU01	R	Desempenho inicial	Todos	"Acréscimo o compromisso..." - benchmark 1	Tempo de execução na primeira tentativa	15s	30 s	10s	20s
2	RU02	R	Desempenho inicial	Todos	"Apague o compromisso..." - benchmark 2	Número de erros na primeira tentativa	0 erro	3 erros	0 erro	1 erro
3	RU03	R	Satisfação - inicial	Secretaria	Questionário: questões...	Média das avaliações	??	7,0	8,5	9,5
4	RU04	M								
5	RU06	R								

9



### Atributo de Usabilidade

- Atributo de usabilidade é uma característica geral de Usabilidade a ser usada como critério para avaliação da interface.
- É essencial já que fornece parâmetros para se medir a usabilidade de uma versão da interface.
- Pode-se usar os cinco principais atributos definidos por Nielsen e outros relevantes.

10



- Na escolha dos atributos deve-se considerar os diversos perfis de atores e as tarefas analisadas.
- Pode ser interessante utilizar-se atributos associados a papéis de usuários específicos.
  - Pode-se usar uma tabela específica para cada papel de usuário focal, por exemplo.

11



- Para uma tarefa complexa, pode ser interessante estabelecer diferentes atributos, por exemplo, aprendizado e desempenho a longo prazo.
- Muitas vezes não é viável estabelecer metas para todas as classes de usuários ou tarefas possíveis.
  - Especula-se que aproximadamente 80% dos usuários usam somente 20% das funcionalidades de um sistema interativo.

12



### Sugestão de atributos

- Desempenho inicial : **refere-se ao desempenho do usuário durante a primeira vez de uso do sistema.**
- Desempenho em prazo longo: **refere-se ao desempenho do usuário em uso mais regular do sistema durante um longo período.**
- *Aprendizagem:* **rapidez e a facilidade com que o usuário aprende a lidar com o sistema.**

13



### Sugestão de atributos

- Retenção: **capacidade do usuário reter na memória o que aprendeu em um período de tempo quando volta a utilizar a interface.**
- Uso de características (features) avançadas: **utilizado para determinar a usabilidade de funções complexas da interface.**
- Primeira impressão: **avaliação inicial do usuário.**
- Satisfação do usuário a longo prazo: **avaliação do usuário após a utilização do sistema por um período maior.**

14



- Prevenção de erros: **capacidade do sistema de evitar erros em sua utilização.**
- Acurácia: **capacidade de oferecer resultados na precisão desejada pelo usuário.**
- Confiabilidade
- Clareza

15



### Instrumento de medida

- **Descreve o método utilizado para se obter valores de um atributo particular de usabilidade.**
- **Deve ser quantitativo, isto é, pode ser medido numericamente.**

16



- A medida pode ser objetiva ou subjetiva.
  - Objetiva - medidas quantitativas do desempenho observável do usuário durante a realização de tarefas usando a interface.
  - Subjetiva - medidas quantitativas baseadas nas opiniões do usuário sobre a interface.

17



- Os termos objetivo e subjetivo referem-se ao modo como são obtidos os dados relacionados com os atributos de usabilidade.

18



- Medidas objetivas: quando os dados são coletados por observação do desempenho do usuário em tarefas de *benchmark*.
- Medidas subjetivas: quando os dados são obtidos através de questionários de preferências dos usuários.
- Medidas objetivas e subjetivas são igualmente importantes para a especificação e para a avaliação da usabilidade de um desenho.

19



Benchmark



20



### Tarefa de Benchmark

- É uma tarefa usada como referência para as medições objetivas.
- Uma tarefa de *benchmark* deve ser bem redigida para indicar claramente o que se deseja e permitir comparações
  - Ex. “marque um compromisso com o Doutor Pacheco por 3 semanas ...”
- Tarefas de benchmark devem ser específicas para que o participante não desvie sua atenção para detalhes irrelevantes durante o teste.
- As tarefas de benchmark devem ser descritas na linguagem do domínio da aplicação.

21



- Ao definir tarefas de *benchmark* diga o que o usuário deve fazer mas não como deve fazer.
  - Ao iniciar uma tarefa, uma pessoa tem a intenção ou necessidade de realizar uma ação. Ao se definir uma tarefa de benchmark, deve-se ter o cuidado de colocar o usuário nesta situação.
  - O objetivo da avaliação é justamente avaliar como o usuário realiza a tarefa utilizando a interface que, no caso, é o instrumento para sua realização.
  - Ex.: use “adicione um compromisso de almoço com os diretores João e Gilberto todas as quartas as 14:00hs por 3 meses” e não “vá no menu de compromisso, entre na tela de repetição de compromisso e ...”.
- A redação da tarefa deve permitir avaliar se o usuário consegue perceber como realizar a tarefa, medir o tempo que ele gasta para isso e/ou a quantidade de erros que ele comete.

22



- Na escolha de tarefas de benchmark, é bom usar características (*features*) simples da interface ou grupos pequenos de características, para que a causa dos problemas identificados possam ser mais facilmente rastreadas no desenho.
  - Quando as tarefas são complexas, elas devem ser convertidas em sub-tarefas para a avaliação, porque as tarefas complexas são demoradas para se executar e dificultam a identificação de problemas
  - Por outro lado, pode ser necessário usar-se uma sequência de tarefas, como no caso em que elas costumam ocorrer em sequência.
    - Ex.: tarefas de pesquisar e modificar um compromisso.

23



### Questionário

- Questionários podem ser usados para avaliar a **satisfação subjetiva do usuário com a interface.**
- Objetivos: além de utilização em testes de usabilidade, o questionário é um instrumento que pode ser utilizado para:
  - guiar o desenho ou melhorias de desenho da interface;
  - identificar áreas potenciais para introdução de melhorias;
  - Validar avaliações comparativas;

24



- Existem métodos científicos para elaboração e validação de questionários.
  - Um questionário deve relacionar várias características de interface de usuário, de uma forma organizada.
    - Montar um questionário não é somente montar uma lista de características, listá-las e colocá-las juntas.
  - De preferência, profissionais devem ser utilizados para sua produção.
  - Um questionário deve endereçar problemas de confiabilidade e validação, criando uma medida confiável para determinados tipos de interfaces.

25



## QUIS (Questionnaire for User Interface Satisfaction)

- O QUIS (University of Maryland) é um exemplo conhecido de questionário para avaliação de satisfação de usuários.
  - Pode ser obtida mediante licenciamento.
  - Muitas vezes, é interessante adaptar-se o QUIS a uma situação específica.

26



- O QUIS contém um questionário demográfico, uma parte dedicada à medida de satisfação de modo geral e uma parte dedicada à medida de onze fatores específicos de uma forma organizada hierarquicamente.
  - Fatores: aspectos de tela; terminologia e *feedback* do sistema; aprendizado; características do sistema; documentação técnica; tutoriais *on-line*; multimídia; reconhecimento de voz; ambiente virtual; acesso à internet e instalação do software.

27



## Valor a ser medido

- Indica o tipo de medida para o qual os valores dos dados são coletados durante os testes juntos aos participantes.
- Medidas mais comuns:
  - tempo em que se completou uma tarefa;
  - número ou percentagem de erros.
- É necessário definir exatamente o que significa um erro.
  - Por exemplo, se o usuário não usa um botão ou menu esperado na realização de uma tarefa, ainda que ele seja desnecessário, deve ser contado como erro.
  - Isso porque será necessário uma correção

28



- Para um questionário, é tipicamente utilizada a média de avaliações medidas.
- O “valor a ser medido” de um atributo como “primeira impressão” na Tabela ERU pode ser obtido como uma média entre vários itens de um questionário.
- Outro exemplo de “valor a ser medido” que pode ser interessante é a percepção do usuário do tempo decorrido.
  - Ex.: uma instalação demorada mas na qual o usuário fica ocupado trocando disquetes pode ser percebida como rápida.

29



- Algumas outras medidas que podem ser usadas nas tarefas de benchmark:
  - Porcentagem de tarefas completadas em um tempo determinado.
  - Proporção sucesso / fracasso.
  - Tempo gasto em erros e recuperação.
  - Número de comandos/ações usados para realizar uma tarefa.
  - Frequência do uso do help e documentação.
  - Número de repetições ou falhas de comandos.
  - Número de comandos disponíveis não executados.
  - Número de vezes em que o usuário expressou frustração ou satisfação.

30



## Níveis de desempenho

- Na tabela ERU, os níveis de desempenho referem-se a metas quantitativas de usabilidade em uma interface.
- O tempo atribuído a cada tarefa depende de sua complexidade e uso.
  - Por exemplo, em uma tarefa freqüente, a duração admitida deve ser menor.

31



## Fontes/critérios para estimativa de níveis

- Um sistema existente ou versão anterior de um novo sistema sob desenvolvimento.
- Sistemas concorrentes, principalmente aqueles com uma grande fatia do mercado ou com uma interface de usuário reconhecida pela qualidade.
- A realização de tarefas sem o uso de um sistema de computação (ex. manualmente, usando papel e caneta).

32





- O uso pelos desenvolvedores de seu próprio protótipo para alguma versão da interface.
- Feedback de mercado, baseado na aspiração dos usuários com sistemas similares.
- Alguma escala absoluta, quando há pouco com o que se comparar.

33



- Diferentes papéis de usuários podem significar necessidade de diferentes tarefas e diferentes níveis de desempenho nas tarefas.
  - Pode-se inclusive usar diferentes tabelas de especificação de Usabilidade.
- Com a prática, desenvolvedores tornam-se bastante habilidosos para estabelecer especificações de usabilidade confiáveis e estabelecer níveis razoáveis de valores para os atributos.

34



### Nível atual

- O nível atual é o nível corrente do valor a ser medido para o atributo de usabilidade na presente versão do sistema.
- Este nível pode ser utilizado não só quando o sistema já está operacional, mas mesmo quando ele ainda está em desenvolvimento ou se trata de um protótipo.
- A medição do nível atual ajuda a assegurar que os outros níveis possam ser estimados.
- É útil saber como está o nível atual de desempenho em relação a um ou mais sistemas concorrentes.

35



### Exemplo

- Na tabela exemplo, foi atribuído o valor do nível atual de desempenho para “apague o compromisso ...” como 0 já que em uma agenda em papel não é esperado erro para esta tarefa.
- O valor do nível atual para “satisfação inicial” foi considerado não aplicável já que não interessa avaliar-se esse atributo para agenda em papel.

36



Ordem	Identificador	Requisito ou meta?	Atributo de usabilidade	Ator	Instrumento de medida	Valor a ser medido	Níveis de desempenho			
							Atual	Pior aceitável	Melhor Possível	Alvo
1	RU01	R	Desempenho inicial	Todos	"Acrescente o compromisso..." - benchmark 1	Tempo de execução na primeira tentativa	15s	30 s	20s	10s
2	RU02	R	Desempenho inicial	Todos	"Apague o compromisso..." - benchmark 2	Número de erros na primeira tentativa	0 erro	3 erros	1 erro	0 erro
3	RU03	R	Satisfação - inicial	Secretaria	Questionário: questões...	Média das avaliações	??	7,0	9,5	8,5
4	RU04	M								
5	RU06	R								

37



## Pior aceitável

- Indica o pior nível de desempenho do usuário que seria ainda aceitável para cada atributo de usabilidade; não o pior que pode acontecer.
- O pior nível aceitável é o nível mínimo de performance que os usuários podem alcançar e ainda considerar-se que a interface possui algum crédito em usabilidade.
- Para todos os atributos avaliados, o nível "pior aceitável", no mínimo, deve ser alcançado.

38



- Diretrizes para se determinar o nível "pior aceitável"
  - deve, quando possível, estar próximo do valor do nível atual do sistema
  - dever ser mais alto na medida em que o nível atual não seja satisfatório.
  - Como o sistema atual pode ser muito diferente do sistema planejado, como no caso temos agenda em papel x agenda eletrônica, pode acontecer que nível atual seja muito rigoroso para ser o pior nível aceitável.
    - Por exemplo, tarefas simples como "acrescentar compromisso" podem ser feitas muito rapidamente em papel. Sendo assim, o desempenho "alvo" e o "pior aceitável" foram estimados como aquém do nível "atual" (pior do que o nível "atual")

39



## Nível alvo

- Indica o valor alvo que significa sucesso inquestionável de usabilidade para a interface, isto é, o nível "que você deseja".
  - O nível Alvo deve ser alcançado para todos os atributos avaliados.

40



- **Diretrizes para se determinar seu valor.**

- É usualmente mais alto que o nível atual para o sistema/versão existente, para representar melhoria.
- Comparar com sistemas concorrentes.
- Se o nível alvo planejado for alcançado durante teste com o usuário, podemos estar confiantes de boa qualidade em termos de usabilidade.

41



### **Melhor possível**

- Indica o limite superior realístico do estado de arte, o “nível de inspiração” de um atributo de usabilidade.
- Mostra o potencial de um atributo e serve como referência para futuras versões do sistema.
- Deve que ser viável, ainda que difícil, atingi-lo.

42



- O melhor nível possível é o melhor nível de performance que você pode esperar para uma situação ideal, o que nos leva à seguinte diretriz:

- considerar onde é possível se chegar com os melhores usuários (mais bem treinados), nas melhores condições, com o melhor desenho e com o melhor uso da tecnologia disponível.

43



### **Resultados observados**

- Os resultados observados nas avaliações são valores reais obtidos observando-se os usuários durante a realização dos testes.
  - Pode ser registrado uma média de valores observados
- Os resultados observados permitem uma comparação rápida entre os níveis especificados e o resultado real dos testes com os usuários.

44



### **Diretrizes**

- Cada atributo de usabilidade deve ser mensurável na prática
  - Por exemplo, é razoável fazer-se uma medição de desempenho do usuário por um longo tempo?
- Os papéis-de-usuários aplicáveis devem ser especificados de forma clara
  - Pode-se usar uma coluna de Atores na tabela ERU, como mostrado, ou mesmo fazer-se tabelas separadas para cada classe de usuários.

45



- O número de atributos a ser medido deve ser razoável na prática
  - Quando o desenvolvedor não tem muita experiência, não deve ser muito ambicioso.
  - Considerar os recursos/prazos disponíveis.
- Todos os membros do projeto devem concordar com os atributos e valores na tabela ERU
  - Isso é importante para o comprometimento da equipe.

46



- Verificar se as metas para os vários níveis são razoáveis
  - É comum o desenvolvedor iniciante ser muito leniente, o que não é produtivo.
- Nos teste de usabilidade, quando os resultados observados são muito piores que os planejados há duas possibilidades:
  - o processo está caminhando normalmente mas há sérios problemas de usabilidade na interface que devem ser resolvidos,.
  - Os valores planejados são irrealísticos.


47



- Verificar se os atributos utilizados refletem as prioridades de Usabilidade
  - A escolha de uma tarefa não representativa pode representar investimento em uma função que não será muito utilizada - perda de dinheiro e tempo!


48





### **Exercício: especificação de usabilidade**

- **Objetivo:** ganhar experiência especificação de usabilidade.
- **Tempo:** 30 minutos
- **Atividades:** produza pelo menos 3 especificações de usabilidade para seu trabalho prático. Procure ser preciso ao descrever as tarefas.
- **Resultados:** tabela ERU preenchida.



49