Avaliação de interfaces

## Computadores e softwares desenvolvidos presumidos que os usuários deveriam ter conhecimento e formação para opera-los:

* + Vocabulário técnico
  + Arquitetura de computadores
  + Resolver problemas computacionais
* Popularização dos computadores
  + Insatisfação/frustração de usuários não especialistas
* Usabilidade
  + Objetivo para desenvolvimento de softwares interativos voltados para o público em geral

## Avaliação:

* Revisar, experimentar ou testar uma ideia um software um produto ou um serviço
* Permite corrigir problemas relacionados com a qualidade de uso antes de inserir o sistema interativo no cotidiano

Perspectiva de avaliação:

* Desenvolvedores
  + Verificar se o sistema funciona de acordo com as especificações de requisitos
* Analistas/usuários
  + Verificar se o sistema apoia adequadamente dos usuários a atingirem seus objetivos de um contexto de uso

## Métricas:

* Taxa de tempo da tarefa (TTR)
  + Tempo para concluir a tarefa
  + 5 usuários descobrem 80% dos problemas (Jakob Nielsen);
  + TTR = TOTAL DE SEGUNDOS / Nº DE USUARIOS
* Taxa de Sucesso da Tarefa (TSR)
  + Eficácia da solução
  + Porcentagem de tarefas concluídas corretamente
  + TSR = Nº TAREFAS CONCLUIDAS COM SUCESSO / TOTAL DE TENTATIVAS
* Taxa de Erro do Usuário (URS)
  + Número de vezes em que houve uma entrada malsucedida (teclado, mouse, etc.)
  + URS = TOTAL DE ERROS OCORRIDOS DE TODOS OS USUATIOS NA TAREFA / TOTAL DE TENTATIVAS \* Nº DE POSSIVEIS ERROS
* Usabilidade Percebida
  + Avaliação Subjetiva da experiência do usuário ao usar um produto
    - Questionário pós-tarefa
      * Preenchidos imediatamente após a conclusão de uma tarefa
      * Capturam as impressões dos participantes sobre uma tarefa
    - Questionário pos-teste
      * Administrados no final de uma sessão
      * Percepção geral dos usuários sobre a usabilidade do site ou aplicativo
      * Impressões sujeitas ao efeito de pico

Efeito de pico: Partes mais intensas e final da experiência (positiva ou negativa) impactam mais as lembranças e avaliações dos participantes

* Questionário padrão
  + Apoiados por muitas pesquisas podem demonstrar:
    - Validade
      * Realmente medem o que pretendem medir
    - Confiabilidade
      * Usuários responderão consistentemente as perguntas da mesma maneira
    - Sensibilidade
      * Podem detectar diferenças significativas
* System Usability Scale
  + Instrumento de pós-teste
    - Ao final de uma sessão de testes de usabilidade
  + Contem 10 questões de escala Likert
    - Abordam a usabilidade e aprendizagem de um sistema
  + Produz uma pontuação de 0 a 10
* As 10 (copiei 8) perguntas da escala de usabilidade do sistema
  + - Eu acho que gostaria de usar esse sistema com frequência
    - Eu acho que o sistema desnecessariamente complexo
    - Eu acho que o sistema é fácil de usar
    - Eu acho que precisaria de ajuda de uma pessoa com conhecimentos técnicos para usar o sistema
    - Eu acho que as várias funções do sistema estão muito bem integradas
    - Eu acho que o sistema apresenta muita inconsistência
    - Eu imagino que as pessoas aprenderão como usar sistema rapidamente
    - Eu achei o sistema atrapalhado de usar
  + Perguntar impares = frases positivas
  + Perguntas pares = negativas
    - Evitar trapaças
      * Seleciona a mesma resposta para todas as perguntas
    - Neutraliza a tendência natural das pessoas de concordar com as afirmações ao invés de discordar delas
* Calculo de pontuação SUS
  + Somar a pontuação total de todas as questões ímpares
  + Subtrair 5 do total para receber x
  + Somar a pontuação dos pares
  + Subtrair esse total 25 para chegar em x
  + Somar a pontuação total dos novos valores e multiplicar por 2,5
    - ((impares-5) + (25 – pares)) \* 2,5 = resultado%
* Single Ease Quentin
  + Instrumento de pós tarefa
    - Quando uma tarefa acabou de ser concluída
  + Faz uma única pergunta ao usuário
    - Exige pouco tempo e esforço dos participantes
    - É minimamente perturbador
  + Utilidade
    - Identificar as partes da interface
* Task Load index (nassa-tlX)
  + Questionário pós tarefa
  + Mede a carga de trabalho percebida dos tripulantes aeroespaciais
  + Usado por muitos estudos em fatores humanos e ergonomia
  + Contem 6 perguntas com uma escala rotulada de 21 pontos
    - Variando de muito baixo a muito alto
  + Cada pergunta aborta uma dimensão da carga de trabalho percebida
    - Demanda mental
    - Demanda física
    - Pressão de tempo
    - Sucesso percebido com a tarefa
    - Nível geral de esforço e nível de frustração
* AttrakDiff
  + Mede como os usuários percebem o design visual das interfaces gráficas
  + Permite avaliar como é experimentada a atrativa do produto
    - Em termos de usabilidade e aparência.

## Métodos de Avaliação

Métodos empíricos: Participação de usuários

Métodos não empíricos: realizados por especialistas

* Níveis de severidade
  + 0 – Sem importância – não afeta operação da interface
  + 1 – Cosmético – afeta apenas a aparência da interface
  + 2 – Simples – problema simples de usabilidade
  + 3 – Grave – Dificulta a finalização da tarefa, mas com esforço é possível
  + 5 – Catastrófico – impede o usuário de finalizar a tarefa

## Métodos Empíricos

* Teste de usabilidade
  + - Técnica de pesquisa utilizada para avaliar
      * Produto
      * Serviço site
      * Aplicativo
      * Protótipo
    - Realizado por representantes do público alvo
  + Descoberta de pesquisas
    - Identificar falhas
  + Testes de benchmark
    - Estabelecer um comparativo entre duas versões de uma aplicação
  + Teste competitivo
    - Comparar versões do seu sistema com uma aplicação da concorrência
  + Eye-tracking
    - Rastreamento ocular
    - Investia os movimentos oculares dos participantes
    - Reconhecer as seções e pontos que os usuários dão mais atenção
  + Teste de aprendizado
    - Avalia o quão fácil é aprender a lidar com a interface
  + Planejamento
    - Objetivos do teste
    - Métricas monitoradas
    - Funções e paginas testadas
    - Metodologia de teste
    - Roteiro e organização das etapas
    - Perfil e quantidade de usuários que participarão
  + Termos de consentimento
    - Termo de consentimento livre e esclarecido
    - Termo para uso da imagem e som da voz
  + Observar usuários
    - Thinking alound (pensar em voz alta)
  + Registrar as seções de teste
    - Tempo para realizar cada tarefa
    - Quantidade de tarefas concluídas/não concluídas
    - Quantidade/tipo de erros
    - Inteirar problemas levantados
    - Pensar em soluções para problemas
  + Analise
    - Estudar os dados obtidos e os feedbacks dos usuários
    - Buscar padrões e tendências escondidas no comportamento
  + Exemplo:
    - Posicionamento dos itens mais efetivos
    - Eliminação de elementos visuais desnecessários
    - Texto que precisa ser reescrito para ficar mais convincente, etc.
* Grupo Focal
  + Entrevista em grupo
  + Entre 5 e 10 pessoas
  + Contar com um moderador para liderar a discussão
  + Objetivo:
    - Reunir informações detalhadas sobre um tópico especifico, a partir de um grupo de participantes selecionados
  + Definir o problema a ser avaliado
    - O que será avaliado?
    - Por que será avaliado?
  + Definir o perfil dos usuários que participarão
  + Escolher a equipe
    - Facilitador, lidera os questionamentos
    - Relator, para anotar a discussão

## Métodos não empíricos

* Avaliação Heurística
  + Usado para encontrar erros de usabilidade em uma interface
  + Envolve um grupo de avaliadores
    - No mínimo três
    - No cada avaliador encontra em torno de 35%
    - 3 avaliadores encontram em média 75% dos erros
  + Formulário:
    - Atividade;
    - Problema;
    - Heurística violada: 1 a 12;
    - Severidade: Catastrófico, grande, pequeno, cosmético;
    - Local;
    - Sugestão de melhoria.
* Percurso Cognitivo
  + Percorre caminhos por uma determinada interface buscando se colocar no lugar dos usuários
  + Etapas:
    - Definir o que será testado
    - Descrever os usuários
    - Determinar tarefas e objetivos a serem testados
    - Elaborar fluxos de atividades para chegar nos objetivos